Assessing real-world music listening in concerts

ASSESSING REAL-WORLD MUSIC LISTENING IN CONCERTS:

Aesthetic experiences and peripheral physiological responses

DISSERTATION

to obtain the degree of Doctor at Maastricht University,
on the authority of the Rector Magnificus, Prof. dr. Pamela Habibović in accordance with the decision of the Board of Deans,
to be defended in public
on Monday 2nd of October 2023, at 16:00 hours

by

Anna Maria Czepiel

Supervisors:
Prof. dr. Sonja A. Kotz
Prof. dr. Melanie Wald-Furhmann (Goethe University Frankfurt / Max Planck Institute of Empirical Aesthetic)

Co-supervisor:
Dr. Lauren K. Fink (McMaster University)

Assessment Committee:
Prof. dr. Bernadette M. Jansma (Chair)
Prof. dr. Elvira Brattico (Aarhus University)
Prof. dr. Antonio Camurri (University of Genova)
Dr. Michael Schwartze

Assessing real-world music listening in concerts: aesthetic experiences and peripheral physiological responses

Anna Maria Czepiel

Cover illustration: Magdalena Czepiel

ISBN: 978-94-6469-584-7

https://doi.org/10.26481/dis.20231002ac

The work in this dissertation was conducted at the Max Planck Institute for Empirical Aesthetics, Frankfurt, Germany, and at Maastricht University, Maastricht, The Netherlands.

Table of Contents

Chapter 1

General Introduction

Listening to music can evoke aesthetic experiences (e.g., Brattico & Pearce, 2013; Menon & Levitin, 2005; Salimpoor et al., 2011), especially in live settings such as concerts (Gabrielsson & Wik, 2003; Lamont, 2011; Sloboda, 2010). However, there is currently limited empirical evidence of music listening experiences beyond typical laboratory settings (Brattico, 2021; Tervaniemi, 2023; Wald-Fuhrmann et al., 2021). Listening to clips of music alone in an isolated room is a completely different experience than going with friends to experience a world-class orchestra perform your favourite Mahler symphony. To maintain the advantage of immersive experience of naturalistic settings with minimal disruption of the music listening experience, continuous objective responses can be effectively captured with peripheral physiological responses. This thesis research investigated real-world music listening experiences in audiences attending Western classical concerts, captured by aesthetic and peripheral responses. This chapter introduces first the general background of empirical aesthetics. The importance of naturalistic contexts to test aesthetic responses is highlighted, with one ideal context being the Western classical concert. Within the scope of this thesis, two concert variables were investigated. First, the concert provided an opportunity to assess aesthetic and physiological responses that are influenced by features of full-length naturalistic music. As a key part of the concert experience is seeing a musician perform, the second variable explored was the visual, next to the auditory, input of seeing a musician perform. This introduction further reviews the measurement and the operationalisation of aesthetic experiences and their potential physiological correlates.

1 Theoretical framework of aesthetic experience of music listening

The first record of an empirical study of an aesthetic response to music was likely Pythagoras (around 570 - 500 BC, cf. Crocker, 1963). While philosophical aesthetics started in the 18th century (Baumgarten, 1735; see also Levinson, 2009 for an overview), a more formal study of empirical aesthetics was established in the 19th century by Gustav Theodor Fechner (*Introduction to Aesthetics*, Fechner, 1876), Helmholtz (*On the Sensations of Tone as a Physiological Basis for the Theory of Music*, Helmholtz & Ellis, 1875), and then later Daniel Berlyne (*Aesthetics and psychobiology*, Berlyne, 1971). The field of empirical aesthetics was originally defined by Fechner as the 'scientific investigation of beauty and art that begins by collecting specific facts that can later be used to build up general principles and, eventually, a general system of aesthetics' (cf. Nadal & Vartanian, 2022). The last few decades have seen a surge in 'neuroaesthetics'; a subdiscipline exploring the neurophysiological mechanisms of an aesthetic experience (Skov & Nadal, 2023; Zeki, 1999)

Although several empirical models of aesthetics exist, a consistency among them is that they comprise three main components: 1) the 'inputs' of such an experience, related to the person, stimuli, and context; 2) the mechanisms involved in stimulus processing, and 3) the 'outputs' of an aesthetic response, such as liking (Anglada-Tort & Skov, 2020; Brattico et al., 2013; Hargreaves & North, 2010; Leder et al., 2004; Leder & Nadal, 2014; Pelowski et al., 2016; Schindler et al., 2017; Wald-Fuhrmann et al., 2021). To date, most neuroaesthetic research of music pertains to the second component. The last decade has seen an increasing focus on the cognitive and aesthetic processing of music (Brattico et al., 2010; Müller et al., 2009), especially naturalistic music (Alluri et al., 2012; Burunat et al., 2014; Cheung et al., 2019; Di Liberto et al., 2020; Kern et al., 2022; Lillywhite et al., 2022; Omigie et al., 2021; Williams et al., 2022). For example, studies have gained insight into how humans both separate and integrate streams of naturalistic music (Hausfeld et al., 2018, 2021) as well as cognitively process musical features based on long-term knowledge (such as syntactic rules; see (Brattico et al., 2013). However, questions are emerging regarding the 'inputs', especially naturalistic settings, and how 'outputs' can be measured in such settings.

'Inputs' of an aesthetic experience constitute the sound, person, and context. The sound refers to the music; the person refers to the listener's mood, preferences and cultural schemas; and the context refers to the physical and social environment (Brattico & Pearce, 2013; Juslin, 2013; Thompson et al., 2023; Wald-Fuhrmann et al., 2021). Despite their importance, naturalistic settings in particular remain relatively unexplored (Brattico, 2021; Tervaniemi, 2023). Another challenge is assessing 'outputs' of an aesthetic response. Outputs can include engagement, emotional responses, and appreciation (discussed more in Section 3.1) together with their potential neural and peripheral physiological markers. In addressing these gaps in the field, the current thesis focused on the 'input' of a naturalistic context of a concert (outlined in Section 2) as well as the potential 'outputs' (outlined in Section 3), such as emotion (Chapters 2-3), liking (Chapter 4), engagement (Chapters 2-3, 5), and corresponding physiological responses (Chapters 2-5). Although several concert components afford exciting venues for exploration (Wald-Fuhrmann et al., 2021), two key concert components were investigated more closely in the empirical studies of the thesis, namely full-length naturalistic music (Chapters 2-3) and the visual aspect of watching a musician perform (Chapters 4-5).

2 Inputs

2.1 Naturalistic contexts

Most psychological and neuroscientific research is conducted in laboratories, where individuals are typically tested in a sound-proof booth and stimuli are presented via computers. However, a paradigm shift in cognitive neuroscience has seen studies moving on to natural settings. For example, studies investigating learning behaviour have moved to classrooms (Bevilacqua et al., 2019; Dikker et al., 2017), whereas studies exploring the experience of visual art have moved to museums (Pelowski et al., 2017), cinemas (Kostoulas et al., 2017; Muszynski et al., 2018), and theatres (Ardizzi et al., 2020). Such ecologically valid settings can offer new insights beyond laboratory settings as well as evaluations whether laboratory studies generalise to more ecological settings. The replication of results across settings would suggest robustness of effects, while differences between them would suggest that music listening experiences are context dependent.

Empirical music research has likewise moved beyond the laboratory. Experience sampling method studies have found that most naturalistic music listening occurs in everyday settings, i.e., background music listening that accompanies activities (e.g., chores, Juslin et al., 2008; Krause et al., 2015; North et al., 2004; Randall & Rickard, 2017; Sloboda & O'Neill, 2001). Additionally, there is a more 'special' reception of music. 'Special' is defined here as 'environments in which music takes on a "heavier" social or cultural weight than normal' (Sloboda, 2010, p.495), for example, in live concerts (Sloboda, 2010). This 'special' type of music listening was more closely investigated in this thesis.

The rise of the digital age has increased availability and affordability of recorded and live-streamed music (Thompson, 2006). Nonetheless, concerts are still a central part of Western cultures. For example, thousands of people queued online undeterred for hours and spent several hundred dollars for tickets to see Taylor Swift live in concert (Eras Tour, 2023), regardless of the website repeatedly crashing (Sherman, 2022). Although research suggests a plateauing attendance of Western classical concerts (Dearn & Pitts, 2017; Kolb, 2001), people remain motivated to attend classical concerts to engage with a live performance that is unique, special (Brown & Knox, 2017), and emotionally moving (Roose, 2008) as well as to feel a sense of belonging and community (Pitts & Spencer, 2008). It was suggested that an 'art' object is perceived to be of higher artistic value if it is experienced in a 'special' context (Leder et al., 2004; Leder & Nadal, 2014). Well-known musical examples that demonstrate this are the two contrasting performances by the famous violinist Joshua Bell. Audiences were willing to pay 100 dollars for a seat to see him perform in Boston Symphony Hall; whereas at a metro station, the same performance yielded 32 dollars, despite thousands of people passing by the same violinist (Weingarten, 2007). Thus, an aesthetic experience of music may be enhanced if heard in a special setting, such as a concert (hall).

To study a more genuine aesthetic responses in music listening (Gabrielsson & Wik, 2003; Lamont, 2011), studies have moved to pop/rock (Dotov & Trainor, 2021; Swarbrick et al., 2019) and classical music concert venues (Bernardi et al., 2017; Chabin et al., 2022; Coutinho & Scherer, 2017; Phillips et al., 2020; Tervaniemi et al., 2021) as well as opera houses (Balteș & Miu, 2014; Scherer et al., 2019a; for a commentary on this, see Cupchik, 2019; Jacobsen, 2019; Scherer et al., 2019b). Such studies have begun exploring musical emotion (Balteș & Miu, 2014; Coutinho & Scherer, 2017b; Scherer et al., 2019a) as well as aspects of liveness and modality (Coutinho & Scherer, 2017b; Swarbrick et al., 2019; Tervaniemi et al., 2021) in concert experiences. However, in a setting with a plethora of factors, there are still several open questions regarding emotional and aesthetic in music listening. One pertains to the music itself, in particular the experience of full-length musical pieces. The influence of musical features on a musical aesthetic experience might be more practically explored in a concert setting. Indeed, an advantage of conducting studies during Western Classic concerts is the focus on the music itself, an etiquette that was established around the 19th century and is still in place today (Weber, 2001). During concerts, lights in the audience are dimmed and communication/noise from the audience (e.g., excessive coughing) is frowned upon. Rather than being restricted to short clips of music, which is often the case in laboratory studies, this exclusive focus on music in a concert setting provides an optimal place to explore an aesthetic experience in full-length, naturalistic music that unfolds over several minutes or even hours.

2.2 Naturalistic music

'Artificial' stimuli are broadly defined as purposefully created in a controlled manner for an experiment, for example, a sequence of beeps at a particular frequency with some beeps played at a different frequency (i.e., Oddball paradigm). On the other side of this spectrum, 'naturalistic' music is defined as real, non-manipulated music created by a composer and performed by musicians to be appreciated as musical art, e.g., a full 1-hour and 26-minute recording of Gustav Mahler's Second Symphony by Simon Rattle and the Berlin Philharmonic Orchestra. Stimuli can lie somewhere along the spectrum, for example, real classical music recordings that have been manipulated to create dissonant versions (e.g., Sammler et al., 2007). Although limited in external validity, studies using artificial stimuli have allowed drawing clear conclusions on how we respond to specific music features. Such research has built a foundation for better understanding auditory cognition and has allowed research to increasingly use richer naturalistic stimuli, such as spontaneous storytelling, speeches, podcasts (Goldstein, 1980; Silbert et al., 2014; for a review, see Alday, 2019) and full-length musical pieces (Abrams et al., 2013; Alluri et al., 2012; Kaneshiro et al., 2020).

Controlled stimuli are manipulated to evaluate behavioural, physiological, or neural responses; naturalistic music is not. Therefore, to understand how music (features) influences such responses, it is important to characterise music. Audio characterisation can be implemented using score-based, symbolic-based[1], and/or acoustic-based analyses (Müller, 2015, for visualisation see Figure 1). Score-based methods characterise music through theoretically informed analyses of the musical score like identifying harmonic changes and formal structures. Symbolic- and acoustic-based methods can extract music features computationally, commonly referred to as Music Information Retrieval (MIR). Symbolic-based algorithms take machine-readable data formats, for example chord names or Musical Instrument Digital Interface (MIDI)[2] formats (cf. (Müller, 2015). Symbolic-based algorithms can extract features related to human perception, for example harmonic surprise (e.g., **I**nformation **Dy**namics **O**f **M**usic, IDyOM Pearce, 2005; **P**rediction by **P**artial **M**atching-Decay, PPM-Decay model, Harrison et al., 2020). While previous score- and symbol-based analyses extract from discrete notes values, such as pitch and length, acoustic-based analyses assess the sound interpretation, that is, directly extract features from a sound signal. Several tools exist for acoustic-based analysis in a number of programming languages (see https://www.ismir.net/resources/software-tools/; Moffat et al., 2015). Currently, one leading option is the MIRToolbox in MatLab (Lartillot & Toiviainen, 2007). This toolbox offers extraction of 'lower level' features based on the acoustic signal alone, e.g., the root mean square (RMS) function can assess a signal's 'energy', representing loudness (page 81). Other functions are based on low-level models of human perception, for example, the function to extract roughness is based on Plomp and Levelt (1965) and Sethares (1998) (page 139-140). Finally, some functions are based on cognitive models, for example the function to extract the clarity of a tonal key assesses correlations between a window of the audio to tonal profiles (Krumhansl, 1990). The (combination of such) extracted features have good similarity with human perception of corresponding features (Alluri et al., 2012). However, it seems that low-level features (e.g., pitch, loudness) have higher consistency compared to higher level features (e.g., emotion,

---

[1] Of note is that while the musical score is technically also symbolic, this score/note-based and symbolic differentiation takes place to describe audio processing techniques, where score-based methods are carried out manual, symbolic, and acoustic refer to computational methods (Müller, 2015).

[2] A symbolic-based music format, where note pitches, onsets, offsets are coded with numbers.

Lange & Frieler, 2018). Despite certain shortcomings, the MIRToolbox remains a useful tool in extracting features and informing which features might result in certain responses, though care should be taken when extracting and interpreting higher-level features relating to more cognitive and emotional models.

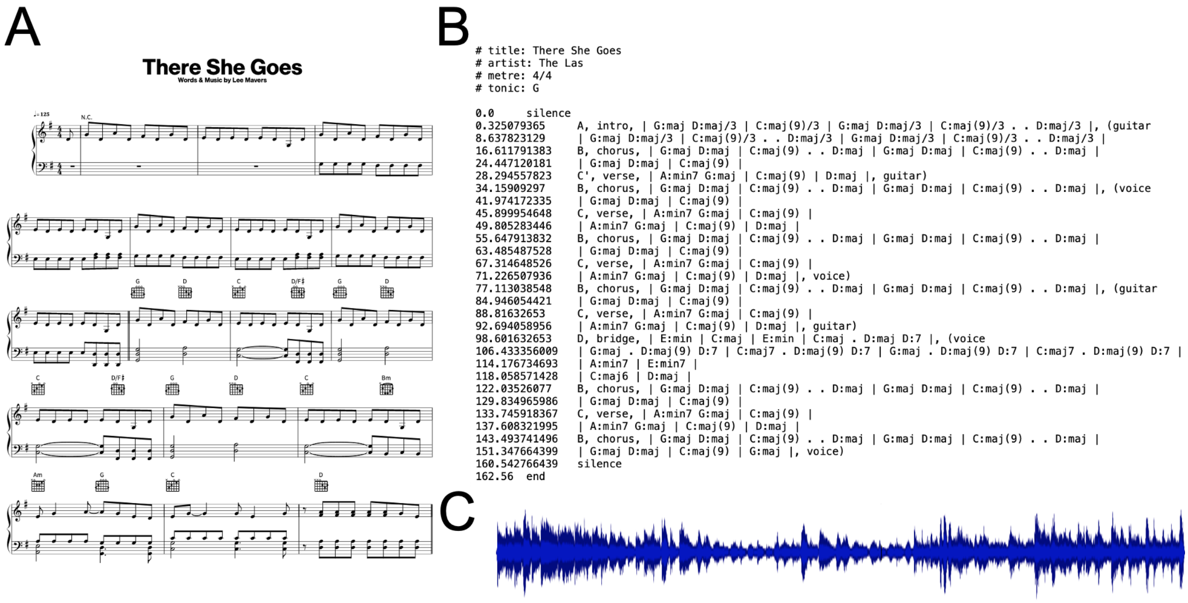


Figure 1. Visual representations of examples of score (A), symbolic (B), and acoustic (C) representations of music. The music is a pop song 'There She Goes' from The La's. B shows one kind of symbolic-based format, i.e., chord representations (taken from https://ddmal.music.mcgill.ca/research/The_McGill_Billboard_Project_(Chord_Analysis_Dataset)/)

Using these kinds of music characterisation techniques, research has linked specific music features to stimulus tracking, emotion, and perception of beauty. With score-based methods, Sloboda (1991) identified that intense emotional sensations seem to be evoked by passages of musical sequences[3], appoggiaturas[4], and new or unexpected harmonies. The last few years have seen an increasing shift to more computational methods, where most studies also characterise music either with symbolic or acoustic-based features. A symbolic-based model that has received a lot of attention is the IDyOM, which extracts values of musical surprise and uncertainty; both of which have been related to neural and physiological correlates of emotional and pleasure responses to music (e.g., Cheung et al., 2019; Di

[3] In this case, a musical sequence refers to a melodic pattern that is repeated at a new pitch level (Drabkin, 2001)

[4] A type of melodic ornamentation in music, where a note that is one step above or below the main (expected note is momentarily sounded, creates a dissonance with prevailing harmony before falling or rising to the main consonant note thereby resolving the harmony (Grove Music Online, 2001).

Liberto et al., 2020; Egermann & McAdams, 2013; Kern et al., 2022). Using acoustic-based methods to assess stimulus 'tracking', the amplitude envelope has been a popular candidate (Di Liberto et al., 2020; Kaneshiro et al., 2020), though recent work shows that spectral flux may be a better overall feature to capture the essence of a stimulus (Schultz et al., 2021; Weineck et al., 2022).

Using rich naturalistic stimuli also allows exploring of several features in parallel. Using acoustic signal-based methods to extract music features, Omigie et al. (2021) reported three subtypes of 'beautiful' music. Low-Tension/Low Energy beautiful music was characterised by lower tempo and polyphony[5]; Low-Tension/High-Energy beautiful music by dynamic changes, major mode and reduced harmonic ambiguity; and High-Tension/High-Energy beautiful music by higher loudness (RMS) and pitch increases. Another study by Alluri et al. (2012) looked at how several music features were processed when listening to a naturalistic piece of a Tango (Piazzolla). Twenty-five features were extracted with the MIRToolbox and underwent dimension reduction (principal component analysis) to obtain more representative characterisation of the music. Timbral features were associated with perceptual and default mode areas of cerebrum as well as cognitive areas of the cerebellum, while tonal and rhythmic features were associated with cognitive, somatomotor, cerebrocortical, and subcortical emotion-related areas. This interaction of cognitive, motoric, and limbic brain regions most likely reflects a clearer picture of how we process 'real-life' music. Despite these advances to better understand emotional and aesthetic responses to naturalistic music listening, this has been scarcely explored in settings that allow further opportunity for more intense aesthetic responses (Gabrielsson & Wik, 2003; Lamont, 2011; Sloboda, 2010).

One key open question about a naturalistic music listening experience in concert settings regards how we experience musical pieces over extended periods of time. So far, most music feature-based analysis has been related to relatively short excerpts of music. In the literature, the longest music pieces are around nine minutes (Abrams et al., 2013; Alluri et al., 2012; Grewe et al., 2007; Kaneshiro et al., 2020). However, there are many (Western classical) music pieces that are longer. In particular, some musical works are made up of multiple individual movements, referred to as multi-movement works (Sadie, 2001). For example, a symphony may have three or four movements (Larue et al., 2001). Individual single movements are self-contained and can be appreciated on their own. Indeed, single movements have been presented to participants in previous studies (e.g., Grewe et al., 2007; Kaneshiro et al., 2020).

[5] Music that has several lines that move to some extent independently from each other (Frobenius et al., 2001; see also Hausfeld et al., 2018).

Nonetheless, it is musically rewarding that all movements from one multi-movement piece are heard in their entirety, as a multi-movement work typically presents contrasting characters that provide interest and enjoyment to a listener. To my knowledge, the only research that explored music listening experience in an entire, multi-movement work has been in opera (Balteș & Miu, 2014; Scherer et al., 2019a). However, the emotional responses to opera are potentially confounded by the dramatic narrative and text. Therefore, it is unclear if they are purely in response to the music or the semantic storyline. Therefore, further work in instrumental music may advance a better understanding of emotional responses to music.

An advantage of naturalistic music is that several features can be explored in parallel (Alluri et al., 2012; Omigie et al., 2021). To date, this has only been explored in laboratory studies. Concert studies have mainly related music listening experience to one musical feature only. Egermann et al. (2013) reported physiological responses related to harmonic expectancy in a monophonic music line. Tervaniemi et al. (2021) found that theta responses increased as a function of improvised music, though it was outside the scope of this study to assess which specific music features influenced this. Therefore, it is of interest to extend these findings not only to richer, polyphonic music, but also explore a wider range of features that characterise such music. Therefore, in addition to characterising the music with just acoustic-based analysis, music was also characterised using both score-based and acoustic-based analyses (Chapter 3).

2.3 Visual input in a concert hall

Visual input is an enriching part of concert performances. The style and design of the concert venue (Coutinho & Scherer, 2017b) or the performer's outfit (Griffiths, 2010, 2011) may also influence the music listening experience. However, the majority of research has focused on the role of musicians and their movements. Visual input of seeing musician movements might increase aesthetic appreciation for a number of reasons. First, compared to unisensory stimuli, multisensory stimuli tend to facilitate perceptual processing as shown by faster response times (Molholm et al., 2002) and neural responses (Arnal et al., 2009; Jessen & Kotz, 2011; van Wassenhove et al., 2005). This ease of perceptual processing may induce more positive feelings in response to a musical piece (Huron, 2012).

Second, visual information in music performances enhances how we perceive musical expression and emotion. In one of the earliest studies on cross-modal music perception, Davidson (1993) found that expressivity cues were mostly captured in the visual domain. Indeed, several subsequent studies compared audio-only (AO), audio-visual (AV), and visual only (VO) versions of music performances, showing that performer movement enhanced a listener's perception of performance quality (Griffiths & Reay, 2018; Tsay, 2013; Waddell & Williamon, 2017). Emotional expression is also stronger when viewing a conductor's (Luck et al., 2010), pianist's (Camurri et al., 2004; Timmers et al., 2006; Vuoskoski et al., 2014), violinist's (Van Zijl & Luck, 2013), clarinettist's (Vines et al., 2004, 2011), percussionist's (Broughton & Stevens, 2009; Dahl & Friberg, 2007), or singer's movements (Lange et al., 2022). As expressive and emotional aspects more easily create more interesting performances (Broughton & Stevens, 2009), this suggests that observing a musician's movement can lead to an enhanced appreciation of the music (see also Platz & Kopiez, 2012 for a meta-analysis).

A further account of why observing movement might enhance an aesthetic experience, is the concept of embodied aesthetics. Stemming from visual neuroaesthetics, Freedberg & Gallese (2007) proposed that seeing (implied) motion in dance/art, evokes embodied simulation of observed movements. This embodied simulation increases – through empathy and emotional contagion – involvement and understanding of the artwork. Such an account suggests that motor simulation enhances aesthetic responses such as liking while observing an art piece that entails (implied) sensorimotor states. This idea has been supported in neuroaesthetic dance studies. Dance movements rated as more aesthetically pleasing evoked increased activation in sensorimotor areas of the action observation network (Cross, 2011) and increased activity of the zygomaticus major (smiling) muscle (Kirsch et al., 2016). When comparing dance sequences where participants had previously a) just heard music of the dances, b) just watched the dance sequences (visual only), or c) watched and rehearsed the dances, the latter two conditions evoked the highest enjoyment ratings (Kirsch et al., 2013, 2015). This motor involvement strongly suggests that motor familiarity increased the pleasure of watching dance sequences. More recently, Finisguerra et al. (2021) assessed motor simulation when participants rated how much they liked an artwork. They found a relationship between motor simulation and empathy in aesthetic experiences. This research advocates an embodied aesthetic experience (Cross, 2015). In music, it is known that rhythmic features in particular evoke activity in motor regions (Grahn & Brett, 2007; Kasdan et al., 2022; Kotz et al., 2018; Zatorre et al., 2007) as well as spontaneous movement (Burger et al., 2013; Gonzalez-Sanchez et al., 2018; Toiviainen et al., 2010) in live music (Swarbrick et al., 2019). Seeing the movements of a musician playing also evoke motor activity (motor cortex, basal ganglia) (Lillywhite et al., 2022). Consequently, the role of the motor system in the aesthetic experience of music is important (Brattico, 2021; Nieminen et al., 2011).

This literature suggests that visual input can enhance art appreciation (Platz & Kopiez, 2012), potentially by means of embodied mechanisms (Freedberg & Gallese, 2007). Although mostly explored in laboratory settings, visual input is beginning to be more investigated in concert settings. Dotov et al. (2020) explored visual social cues in pop concerts. People listened to groovy music in eyes open and closed conditions; the eyes open conditions allowed for an enhanced social music experience. Coutinho and Scherer (2017b) also compared modality in live performances. However, a potential confound in this latter study was the location itself; audio-visual performances were presented in a church setting, while the audio-only stimulation was presented as recordings in a lecture hall. Thus, lower emotional ratings in AO compared to AV could have resulted from the context rather than the modality differences. Additionally, the idea of embodied mechanisms in music has not yet been explored as much as it has been in dance (Cross, 2015). Therefore, questions remain about how AV and AO music performances compare, which are further investigated in **Chapters 4-5**.

3 Outputs

Within the empirical literature, the definition and operationalisation of aesthetic experience or response has been inconsistent. Musical aesthetic experience was operationalised as the perception of beauty, an overall evaluation, or any intense response to art, especially pleasure (Anglada-Tort & Skov, 2020; Hargreaves & North, 2010), 'arousal potential' (Berlyne, 1971), states of aesthetic awe, being moved, experiencing thrills (Konečni, 2005) or 'peak experiences' (Gabrielsson & Wik, 2003; Maslow, 1961). In this thesis, the aesthetic experience of music is related to the following concepts: emotions, aesthetic liking, and engagement.

3.1 Behavioural/self-reports

3.1.1 Music emotion

One key aspect of an aesthetic experience is an emotional response (Brattico & Pearce, 2013). Emotion can be 'perceived' or 'felt'. In music psychology the former refers to the emotion that music communicates, the latter refers to the emotion that music evokes in a listener (Gabrielsson, 2002). Felt and perceived emotion might be the same, for example, if music expresses happiness, which subsequently makes the listener feel happy. However, they might be different, for example, if music is perceived as sad but evokes a listener to be happy (perhaps because the music is beautiful). In felt emotion, it has been controversial whether or not aesthetic emotions exist, based on the (lack of) distinction between such emotions at a psychological and neural level (Koelsch, 2010; Menninghaus et al., 2019, 2020; Scherer & Zentner, 2001; Skov & Nadal, 2020). A key argument for aesthetic emotions is that they are somehow separate from so-called ‘everyday’ emotions and can be related to liking and pleasure. For example, hearing a sad news story might make you sad and upset (everyday emotion), while hearing sad music might make you sad, but also give you pleasure from being moved (aesthetic emotion, Eerola et al., 2018). Additionally, emotions may differ by what evokes them. Aesthetic objects and their contexts such as artworks and concert halls evoke emotions untypical for everyday life, such as awe, being moved, and thrills (Konečni, 2005; Nadal & Vartanian, 2022) as well as mixed emotions like nostalgia (Lahdelma & Eerola, 2015). Indeed, some emotions are only evoked when listening to music in a specific context such as a concert setting (Brattico & Pearce, 2013). Therefore, as an ‘output’ of an aesthetic experience, we measured felt emotions that also encapsulate so-called aesthetic emotions, as the concert setting affords such responses to music.

A key aspect of the current thesis was measuring music-evoked felt emotion. Two main models are used for music-evoked emotion: discrete (i.e., emotional categories) and dimensional models. The main discrete emotion model is the Basic Emotions theory (Ekman, 1992) that identified primary emotions as happiness, sadness, anger, fear, surprise, and disgust. Music-specific emotional categories have also been developed, such as adjective circles (Hevner, 1936; Schubert, 2003) and the **G**eneva **E**motion **M**usic **S**cale (GEMS, Zentner et al., 2008), which later developed into the **Ge**neva **M**usic **I**nduced **A**ffect **C**hecklist (Coutinho & Scherer, 2017a). The most common dimensional model used in music is the circumplex model of affect, which postulates that emotions can be mapped onto two orthogonal dimensions of arousal or valence (Russell, 1980). Research has shown that when we perceive emotions, we classify them both categorical and dimensional, but at different timescales (Giordano et al., 2021). In comparing the discrete and dimensional models for felt musical emotion, ratings were more consistent and precise in the dimensional (arousal-valence) model, compared to the discrete (i.e., basic emotions) and the GEMS model (Vuoskoski & Eerola, 2011). Although the GEMIAC has not been compared to either model (being released in 2017), it is a fitting combination of both discrete (i.e., emotion categories) and dimensional levels (i.e., intensity/frequency of emotion), providing a good solution for assessing felt emotions. This was the approach taken in the data analyses presented in Chapters 2-3.

The mechanisms that might evoke and convey emotions in felt music have received a lot of attention (for reviews see Koelsch, 2010; Vuust et al., 2022). With regard to felt emotion, early research suggested that musical consonance and dissonance (Blood et al., 1999; Helmholtz & Ellis, 1875), and tension and resolution within music (Huron, 2006) evoke emotions. More recently, more holistic frameworks, encapsulating a combination of mechanisms, have been put forward. Scherer and Zentner (2001) proposed a set of mechanisms that underlie evoked emotions in music listening, namely 1) appraisal, 2) memory, 3) empathy, 4) proprioceptive feedback, and 5) facilitating the expression of pre-existing emotions. Another theory has been put forward by Juslin and Västfjäll (2008), with an updated version by Juslin (2013), the '**BRECVEMA**': **B**rain Stem Reflex, **R**hythmic Entrainment, **E**valuative Condition, Emotional **C**ontagion, **V**isual Imagery, **E**pisodic Memory, **M**usical Expectancy, **A**esthetic Judgement. There are two common themes across these theories. First, the 'mirroring' of the emotion expressed in music by the listener via emotional contagion. For example, high arousal in a listener might be evoked by arousing music characteristics, such as fast tempi, loudness, timbre with higher harmonics, and larger pitch variation (Gabrielsson, 2012; Hevner, 1936). Second, the anticipation and expectancy in music can lead to emotions as well as pleasure. Although the idea that music-derived pleasure arises from expectation has been around for several years (Huron, 2006), the predictive processing theory is a more recent proposition. This theory postulates that music listening is an active process, where continuous generation and improvement of musical expectations occur; reward is elicited when we improve certainty of expected notes (Koelsch et al., 2019). This theory is indeed gaining evidence that expectancy and surprise drive musical pleasure in lab studies using MIDI chords (Cheung et al., 2019) and in concert settings using monophonic music (Egermann et al., 2013). However, this is yet to be explored in naturalistic, polyphonic music.

3.1.2 Aesthetic judgements and liking

Aesthetic judgements of music have been related to the perception of beauty in music, particularly in sad music (Eerola et al., 2016; Juslin, 2013; Sachs et al., 2015). Liking is also an evaluative component of an aesthetic experience (Brattico & Pearce, 2013; Hargreaves & North, 2010). In a similar way to emotion, liking can be related to musical features. It has been shown that musical complexity has an inverted U-shape relation to liking; very simple music and very complex music typically receive low liking ratings, whereas music with average complexity has high liking ratings (North & Hargreaves, 1995). Other research shows how emotion may be a predictor of liking. Schubert (2007), for example, found that the more we are moved by music, the more we may like it. Other features such as familiarity have also been linked to liking (Madison & Schiölde, 2017; van den Bosch et al., 2013). As opposed to emotion, liking in music can be more simply operationalised. The most common operationalisation is listeners simply indicating on a rating scale how they like or dislike music.

3.1.3 Music engagement and absorption

Music can induce several kinds of states, such as mind-wandering (Taruffi et al., 2017), trance (Sacks, 2006), and absorption states (Vroegh, 2019), where the latter is a heightened form of engaged listening. Although music engagement can refer to the actual involvement of musical activities such as playing and performing music (Chin & Rickard, 2012), this thesis refers to musical engagement as becoming absorbed in music in real-time (Lamont, 2011, 2012; Olsen et al., 2014), the feeling of being compelled, drawn into what is happening (structurally) in the music, and interest in what will happen next (Chamorro-Premuzic & Furnham, 2007; Schubert et al., 2013). In terms of the aesthetic experience, it was shown that state absorption can have a positive effect on one's enjoyment (Hall et al., 2016; Thompson, 2006) and liking of music (Hogue et al., 2016; Presicce & Bailes, 2019). Musical reward sensitivity has also been linked to the tendency for absorption states in music listening (Cardona et al., 2022).

Engagement in music listening has been mainly measured in laboratory settings. However, rating absorption might be limited in such a setting. Within a concert context, absorption is likely higher compared to a laboratory study, as the setting affords more focus on the music. Indeed, boredom does increase in a more lab-like setting (Coutinho & Scherer, 2017b). Audiences who had higher emotional engagement with music in a concert, seemed to gain more from such experiences (Thompson, 2006). Therefore, it is important to assess such engagement in a setting that allows genuine absorption and more intense emotional responses (Lamont, 2011).

3.2 Biological (peripheral physiology) responses

Music listening is a dynamic experience, in which one's emotional and absorption states are constantly changing. Although participants can continuously self-report their experiences (Egermann et al., 2013; Isik & Vessel, 2019; McAdams et al., 2004), continuous rating is 1) a difficult task and 2) distracts from the actual listening experience. Therefore, continuous measurement of the neural and peripheral responses is an insightful alternative to assess one's experience in an undisturbed way. Exciting new steps have emerged from mobile EEG and peripheral measurement. For example, Chabin et al. (2022) reported theta-related brain synchrony with synchrony of self-reported pleasure in concert audiences watching extracts of orchestral and choral pieces. When comparing brain responses to regular and improvised performances of Western classical pieces (Johann Sebastian Bach and Erkki Melartin), Tervaniemi et al. (2021) found that theta brain activity increased in improvised performances.

Exploration of peripheral responses is also of great interest as there has been recent surge toward exploring relationships between the brain and the autonomic systems such as the heart and respiration (Azzalini et al., 2019; Criscuolo et al., 2022; Klimesch, 2018; Madsen & Parra, 2022; Parviainen et al., 2022). The peripheral system constitutes autonomic and somatic systems, which this section outlines, and reviews how they have been related to emotion, liking, and engagement.

3.2.1 Autonomic physiology

The autonomic nervous system (ANS) constitutes the sympathetic nervous system (SNS) and parasympathetic nervous system (PNS) that represent 'fight-or-flight' and 'rest-and-digest' systems, respectively. Although these systems primarily maintain homeostasis, SNS and PNS responses have been related to high arousal (e.g., happiness, surprise) and low arousal emotions (e.g., calmness, sadness), respectively. Of interest in the current thesis were ANS measures of skin conductance, heart rate, and respiration.

3.2.1.1 Skin conductance

One innervation of the SNS is to the sweat glands; thus, an increase in sweat secretion is an indication of SNS activation (Dawson et al., 2016). Sweat secretion is commonly measured by skin conductance (SC)[6], induced by passing a small current through two electrodes that are placed on the skin surfaces (Dawson et al., 2016; Fowles et al., 1981). If the voltage is held constant between these two electrodes, one can measure the skin conductance (i.e., current flow) in microSimens (μS). An increase of SC reliably indicates short-term arousal, such as an orientation response (OR, Barry & Sokolov, 1993). The skin conductance signal can be decomposed into two components: a slower moving tonic (skin conductance level, SCL) and a faster, event-related phasic response (skin conductance response, SCR) (e.g., Benedek & Kaernbach, 2010). In event-related paradigms, the SCR is typically preferred and has a classic sharp rise and slightly slower fall trajectory (see Figure 2, purple line that represents SCR).

3.2.1.2 Cardiac activity

The heart reflects both SNS and PNS activity. Stimulation of the SNS via the sympathetic nerve increases the rate of heartbeats, while stimulation from the PNS via the vagal (cranial nerve X) nerve decreases the rate of heart beats as well as maintains the heartbeat (of a healthy adult) at around 60-80 bpm (Berntson

[6] Also known as electrodermal activity (EDA) or galvanic skin response (GSR).

et al., 2016; Gordan et al., 2015). Cardiac activity can be recorded by means of an electrocardiogram (ECG, also EKG), electrodes placed on the skin surface, and photoplethysmography (PPG) or blood volume pulse (BVP), measured via a clip placed on a finger or ear that records the amount of red light passing from one side of the clip to the other. Heart rate (HR), measured in inter-beat intervals (IBIs) or beats per minute (bpm), can be measured by calculating time distances between consecutive heart beats, where a 'beat' is related to an R/systolic peak in ECG/BVP signals, respectively (see Figure 2; for a tutorial, see Czepiel, 2022). Increases and decreases of HR reflect higher and lower arousal, respectively. Biphasic deceleration-acceleration changes in HR have been related to the orienting response (OR) and conscious detection and emotional arousal (Bradley, 2009; Graham & Clifton, 1966; Park et al., 2014; Stekelenburg & Boxtel, 2001). The initial deceleration is thought to reflect a preliminary decrease of attention to internal resources, thus an increase in the perception of the external environment, while the subsequent acceleration is thought to reflect internal preparation for action in response to an external stimulus (Stekelenburg & Boxtel, 2001).

HR fluctuations, referred to as heart rate variability (HRV), have been investigated in both the time and frequency domains (Task Force of the European Society of Cardiology, 1996). Time domain HRV measures include the average, range, or measures of dispersion of inter-beat intervals (IBIs), such as the standard deviation of IBIs (also known as NN intervals, SDNN). In the frequency domain, HRV is measured as the power in certain frequency bands. The most common heart rhythms are low frequency (LF, 0.04-0.15 Hz) and high frequency (HF, 0.15-0.40 Hz), though additional frequency bands include very low frequency (VLF, 0.003-0.040 Hz) and ultra-low frequency (ULF < 0.003 Hz). The HF component is thought to reflect PNS activity and correlates negatively with stress (Shaffer & Ginsberg, 2017), but positively with low arousal (Di Bernardi Luft & Bhattacharya, 2015). As HR variations in this frequency band are related to breathing, this frequency is also referred to as the respiratory frequency band (Shaffer & Ginsberg, 2017). LF seems to reflect a combination of both sympathetic and vagal influences; thus, the LF/HF ratio is used to represent SNS activity (Shaffer & Ginsberg, 2017; Task Force of the European Society of Cardiology, 1996).

3.2.1.3 Respiration

Respiration is a third measure of interest that reflects autonomic activity. Similar to the heart, the lungs are also innervated by both SNS and PNS, which dilate and constrict the bronchioles in the lungs, respectively. Similar to the heart, respiration signals are assessed in cycles, with a signal rise and fall reflecting the breathing in and out, respectively. Again, the speed of these cycles – i.e., respiration rate (RR) – is assessed by identifying the 'peaks' of the signal, assessing the difference between these peaks (inter-breath-intervals), and transforming them to units of beats/breaths per minute (bpm) (see Figure 2; for a tutorial, see Czepiel, 2022). Increases in respiration rate (RR) have been associated with arousal (Lorig, 2016), such as excited negative, tense (both positive and negative) affective states, while decreases of RR is related to more calming states (see Boiten et al., 1994 for a review). The respiration system is also innervated by the somatic system, i.e., controlled by voluntary muscles, outlined in the following section.

3.2.2 Somatic system

The somatic nervous system controls the skeletal muscles, i.e., voluntary movement. Here, muscles related to facial expressions are of interest as they can convey emotional states of an individual (Ekman et al., 1983). Certain coding systems and video detection software that can objectively measure and code visible facial movement have been developed to operationalize facial movements, for example the Facial Action Coding System (Ekman & Friesen, 1978). However, the measurement of facial muscle movement via surface electromyography (EMG) recordings is a more optimal solution. Many subtle psychological processes are often not reflected by visible action; EMG signals offer a sensitive, reliable analysis of muscle activity (Tassinary et al., 2016). The most common method to measure EMG is using bipolar electrode pairs that are placed directly above a muscle region of interest (Tassinary et al., 2016). The voltage changes that are detected in surface EMG stem from motor unit action potentials, where the electrical activity associated with these events represents a measure of tension, muscular contraction, or movement (Tassinary et al., 2016). The most commonly measured facial muscles are the 'smiling' and the 'frowning' muscles, that is, zygomaticus major and corrugator supercilii muscles, respectively. These two muscles have been mostly related to the valence. Much visual-orientated research shows that EMG activity for the frowning muscle (corrugator) increases for negatively-valenced stimuli, whereas EMG activity for the smiling muscle (zygomaticus) increases when viewing positively valenced stimuli (Bradley & Lang, 2000; Cacioppo et al., 2000; Larsen et al., 2003). Observing dance movements that are rated as 'liked' evoked higher activity in the zygomaticus major muscle in dancers (Kirsch et al., 2016) (though see discussion below for how these measures are related to music).

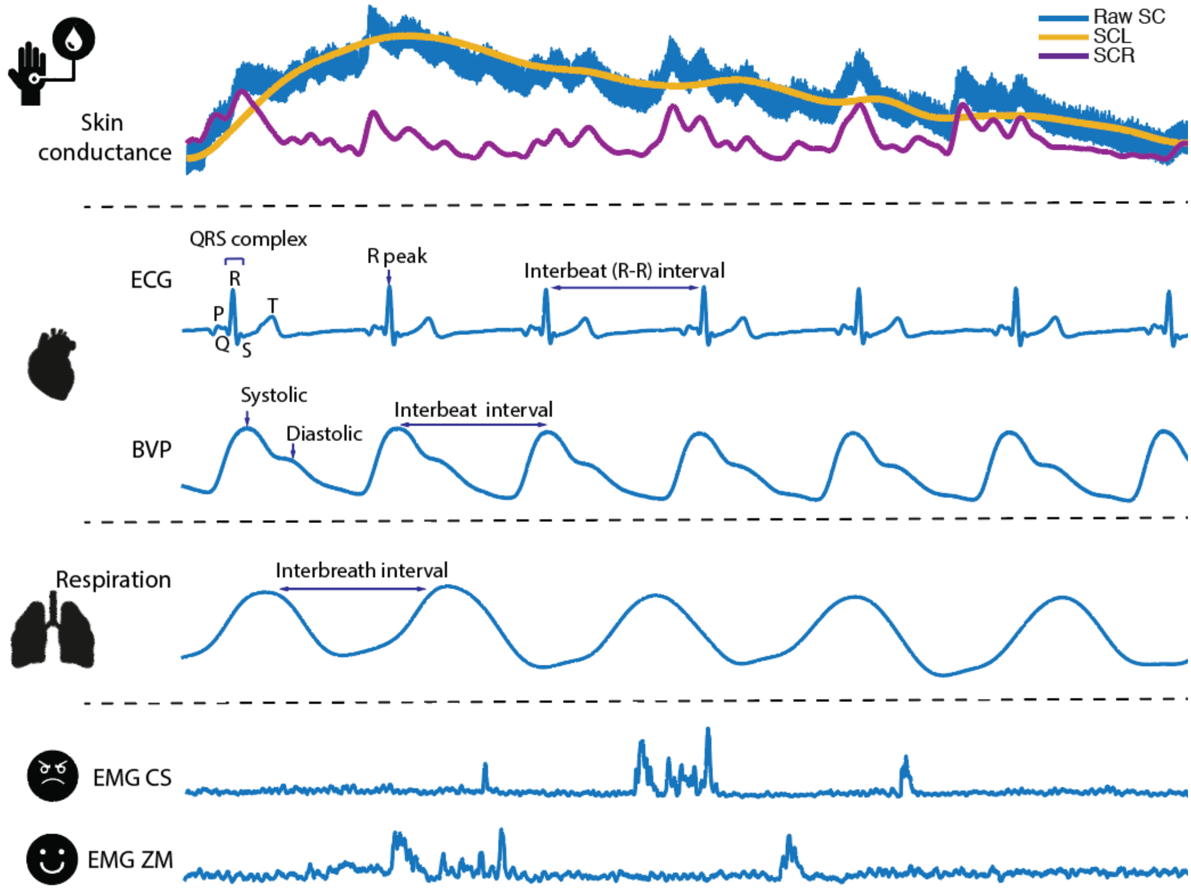


Figure 2. Visual representation of physiological responses. Top panel shows skin conductance (SC, top panel). The raw skin conductance (black line) can be decomposed into components of the slower moving tonic skin conductance level (SCL, yellow line) and the faster moving phasic skin conductance response (SCR, blue) and skin conductance level (yellow). Second panel shows cardiac measures of Blood Volume Pulse (BVP) and Electrocardiogram (ECG). The third panel shows respiration signals. Lowest panel shows examples of facial muscle activity from corrugator supercilii (CS) and the zygomaticus major (ZM) muscle, referred in the literature as the frowning and smiling muscles, respectively.

3.2.3 Synchrony of peripheral signals

While peripheral measures have been used to reflect emotional arousal and valence in music listening, assessing synchrony of peripheral measures has been particularly beneficial in naturalistic paradigms. As synchrony was originally used as a method to enhance signal-to-noise ratio (SNR) using inter-subject correlation, this synchrony measure is outlined next. This and other measures have been likewise linked to engagement (see Figure 3 for an overview and visualisation of all synchrony measures here in this thesis).

3.2.3.1 Inter-subject correlation

Physiological and neural signals can be noisy. Traditionally, paradigms using such measures are trial based, where responses are averaged to assess commonalities across trials to increase the SNR. To enhance SNR in naturalistic (non-trial-based) paradigms, Hasson et al. (2004) pioneered the technique to assess commonalities across several participants, known as inter-subject correlation (ISC). This measure assumes that neural and peripheral responses might reflect: 1) stimulus-triggered processing, which is consistent across participants, 2) individual differences in responses related to timing and intensity, and 3) spontaneous activity unrelated to a stimulus (Nastase et al., 2019). ISC maximises this first component, assuming that if time-locked responses from several participants are highly correlated, there is a high probability that important events occur at these time points in the stimulus that caused a common response in all participants. In other words, the assumption is that it is highly unlikely that noise across several individuals would be correlated (Hasson et al., 2004, 2010; Nastase et al., 2019). To date, there is no single standardised way to calculate ISC. For example, ISC can be calculated in a pairwise or leave-one-out fashion, as well as a global value over time or in a temporal approach (Nastase et al., 2019, 2020). Although originally employed for multivariate fMRI or EEG signals (Abrams et al., 2013; Cohen & Parra, 2016; Dmochowski et al., 2012; Hasson et al., 2004; Kaneshiro et al., 2020), ISC is now more frequently used in univariate signal analysis of peripheral responses such as the heart rate, respiration rate, and skin conductance (Bracken et al., 2014; Golland et al., 2015; Madsen & Parra, 2022; Pérez et al., 2021), making it a promising method for measuring physiological responses to naturalistic music.

3.2.3.2 Inter-subject phase coherence

Next to assessing synchrony of responses between participants in the time domain, i.e., whether signals increased/decrease in a synchronised fashion, synchrony can also be explored in the frequency domain by looking at the phase synchrony between oscillatory response between participants (Kaneshiro et al., 2020; Kauppi, 2010; Sachs et al., 2020). While time-domain synchrony can be assessed with correlation measures, one way to assess phase synchrony is by calculating the coherence between two phase angles (Cohen, 2014). Assessing phase synchrony of heart/respiration cycles would show if participants' heartbeat and breath cycles align (see Figure 3D).

3.2.3.3 Stimulus-Response Correlation

A related way to assess stimulus tracking is to correlate an individual's response to the stimulus. This has been referred to as a Stimulus-Response Correlation (SRC, Dmochowski et al., 2012; Kaneshiro et al., 2020; or SRCorr, Weineck et al., 2022). Here, the acoustic signal is reduced to an overall feature that best represents the auditory stimulus. The envelope and/or its derivative has been a strong candidate as a way to capture overall capture naturalistic speech (Chalas et al., 2023) and music (Di Liberto et al., 2020; Dmochowski et al., 2012; Hausfeld et al., 2021; Kaneshiro et al., 2020; Weineck et al., 2022). However, variation of spectral information, or spectral flux, might represent the how humans perceive acoustic signal more faithfully (Albouy et al., 2020; Obleser et al., 2008, 2012; Obleser & Weisz, 2012). Indeed, spectral flux can account for not only energy, but also spectral information (Müller, 2015) and beat information (Weineck et al., 2022). Compared to the amplitude envelope, the spectral flux was associated with strongest neural synchronisation (Weineck et al., 2022) and evoked peripheral muscle activity (Schultz et al., 2021). Therefore, to characterise the music with one acoustic feature, we extracted spectral flux (Chapter 5). The acoustic feature is then correlated to participant response (see Figure 3A).

3.2.3.4 Stimulus-Response Phase Coherence

As well as assessing synchrony of responses between a participant and an acoustic feature in the time domain, it is also interesting to assess synchrony between a participants' response and an acoustic feature in the phase domain, i.e., the synchronised phase to an external rhythm with coherence measures (Cohen, 2014). This can be implemented in one of two ways. One way is to calculate the coherence between oscillatory responses and actual note onsets. Although studies have shown that respiration aligns to musical beats (Etzel et al., 2006; Haas et al., 1986), there is limited evidence that heartbeats do the same (Ellis & Thayer, 2010; Koelsch & Jäncke, 2015; Mütze et al., 2020). However, this former approach assumes that oscillatory responses can occur at the frequency range of a stimulus. In neural (M/EEG) measures, oscillations in the brain have been found to occur between 1-80 Hz (Klimesch, 2018; Tallon-Baudry, 1999). Therefore, it is possible to assess a range of onsets to specific frequency bands of brain oscillations (Doelling & Poeppel, 2015). However, heart and respiratory oscillations have frequency ranges limited to 1-1.5 Hz (i.e., 60-100 bpm at rest) and 0.2-0.3 Hz (i.e., 12-16 bpm at rest), respectively (though the heart has other rhythms, see 3.2.1.2). Due to this restricted frequency ranges of heart and respiration measures, the approach of transforming acoustic signals to the frequency domain and assessing phase in frequencies in the range of cardiorespiratory rhythms might therefore be preferable.

Such investigations have been conducted in brain oscillations (Chalas et al., 2023; Harding et al., 2019; Kaneshiro et al., 2020; Peelle et al., 2013; Vanden Bosch der Nederlanden et al., 2020); thought questions remain to what extent this could apply also to peripheral rhythms (Criscuolo et al., 2022; Klimesch, 2018; Parviainen et al., 2022).

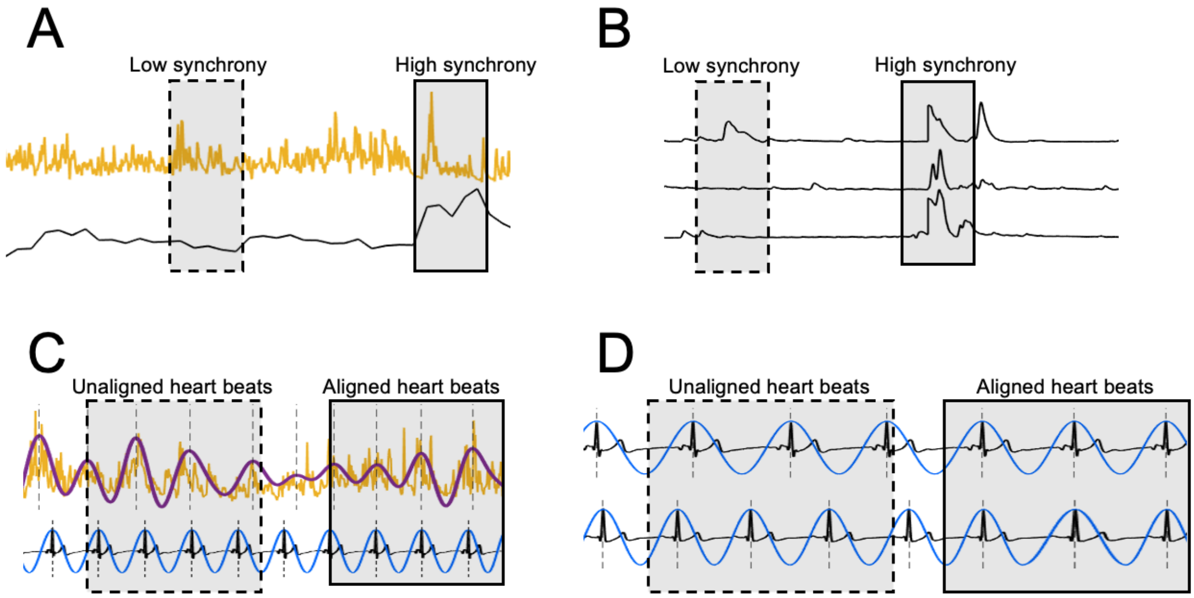


Figure 3. Types of synchrony. **A** represents stimulus-response synchrony, where in the low synchrony case the spectral flux (upper yellow line) increases, while HR (lower black line) decreases. In the high synchrony, both spectral flux and HR increase. **B** represents inter-subject correlation (ISC), showing synchrony of three example participant responses. For low synchrony, there is some increase in the upper participant, but less or no increase in the other participant responses. In high synchrony, all participants have an increase in their skin conductance. **C** shows a representation of stimulus–responses phase coherence, where the unaligned spectral flux and heartbeat would lead to low synchrony. Aligned beats would lead to high phase coherence. **D** represents inter-subject phase coherence (ISPC), which represents how aligned hearts beat are between participants. Unaligned heart beat would yield low phase-domain synchrony, while hearts beat that beat together (i.e., are aligned) would yield high phase-domain synchrony.

3.3 Relating peripheral physiological measures to musical aesthetic experiences

Most music research has linked peripheral physiological responses to emotion, broadly on arousal and valence dimensions. Such responses can be attributed to either mirroring acoustic features of the music itself or reflecting a subjective state in response to the music (Brattico et al., 2013).

In terms of physiological responses that mirror acoustic features, ANS responses have been broadly related to arousal. For example, a sudden change in a stimulus (e.g., Oddball paradigm), have been shown to evoke typical signatures of emotional arousal, such as increased skin conductance and deceleration-acceleration HR patterns. These physiological signatures have been related to an auditory startle reflex (Davis, 1934), orienting response (OR, Barry & Sokolov, 1993), and mismatch negativity (MMN[7])-like reaction (Chuen et al., 2016). On a more cognitive level, physiological arousal responses have been related to harmonic or rhythmic (un)expectancy. For example, SCR tends to increase in response to surprising harmony (Koelsch et al., 2008; Steinbeis et al., 2006) and melodic pitches (Egermann et al., 2013). In designs where physiology is assessed across longer music epochs, the majority of studies show that SCR, HR, RR, and LF/HF ratio power increases with arousing music, while HR decreases and HF power (heart rhythm reflecting PNS activity) increase in response to calmer music (Bernardi et al., 2006; Iwanaga et al., 2005; Krumhansl, 1997; for thorough reviews, see Bartlett, 1996; Hodges, 2009, 2010; Koelsch & Jäncke, 2015). For valence, heart rate responses seem to decelerate more for unpleasant than pleasant music (Bradley & Lang, 2000; Dellacherie et al., 2011; Ruiz-Padial et al., 2011; Sammler et al., 2007). Facial EMG results have shown that activity of the smiling muscle (zygomaticus) increases for positively valenced music, while corrugator activity increased during negatively valenced music (Khalfa et al., 2008; Lundqvist et al., 2008; Roy et al., 2009). However, an increase during typically perceived unpleasant (dissonant) music has also been reported (Dellacherie et al., 2011). Perhaps the activation of the smiling muscle does not necessarily attribute the activity to smiling. Rather, it could represent smiling, a grimace, or ironic laughter. In these prior studies, the physiological response mirrored the emotions expressed in music. Considering the mechanisms underlying emotion induction (see Section 3.1.1), if physiology can reflect the emotion expressed in music, emotional contagion – where the listener feels the perceived emotion of the music – may seem to play a role in the evoked responses.

Physiology has likewise been related to more subjective states. For example, physiological responses have been related to an intensely pleasurable sensation in music listening, namely music-evoked “chills” (frissons) that are characterised by subtle tremor, a slight shudder, or a tingling through the body (Goldstein, 1980; Sloboda, 1991). These chills have been related to increased SCR, HR, and RR (Blood & Zatorre, 2001; Craig, 2005; Grewe et al., 2009; Guhn et al., 2007; Harrison & Loui, 2014; Koelsch & Jäncke, 2015; Salimpoor et al., 2009; Steinbeis et al., 2006). In terms of aesthetic judgments, beautiful music of

[7] Mismatch negativity (MMN) is a negative event-related brain potential component, which traditionally reflects a detection of stimulus change. For example, in a sequence of repeating tones at 1000 ms inter-onset intervals (IOI), a deviant of a 1100 ms IOI could cause a MMN. MMN can be elicited by changes in a range of features, such as intensity, pitch, and timbre (Näätänen & Kreegipuu, 2011),

high energy evoked increases in SC, while low energy beautiful music evoked increased respiration rate and smiling muscle activity (Omigie et al., 2021). Physiological responses have likewise been related to pleasure, where increases of physiological arousal occur at highly pleasurable moments. However, it should also be noted that this increase in physiological arousal does not always lead to a positive experience; Merrill, Ackerman & Czepiel (2023) have also reported high physiological arousal to highly disliked music, especially when listeners felt an unpleasant experience while listening to music.

While physiological responses are related to emotion, the synchrony of peripheral and neural responses have been related to engagement. Several studies have linked stimuli-response and inter-subject synchrony to engagement with auditory stimuli, such as in movies (Cohen & Parra, 2016; Dmochowski et al., 2012; Kang & Wheatley, 2017), speech (Schmälzle et al., 2015) and music (Kaneshiro et al., 2020; Madsen et al., 2019; Pérez et al., 2021). This was linked to the fact that increased ISC was related to emotionally engaging timepoints in stimuli and decreased alpha-band activity (Dmochowski et al., 2012; Kang & Wheatley, 2017), where alpha activity (8-12 Hz) is a neural signature that reflects inhibitory attention states (Klimesch et al., 2007). A decrease in ISC has been related to repeated presentations of familiar music (Madsen et al., 2019) or movies (Dmochowski et al., 2012), suggesting that participants might be more engaged when first listening/watching stimuli, but might disengage during repetitions. In addition to time domain synchrony, studies have also explored phase synchrony in relation to engagement. Neural oscillations have been shown to phase synchronise to external rhythms (Henry & Obleser, 2012; Lakatos et al., 2005), such as speech (Luo & Poeppel, 2007; Peelle et al., 2013) and music (Harding et al., 2019; Vanden Bosch der Nederlanden et al., 2020). Entrainment is beneficial in that it aids processing (Schroeder & Lakatos, 2009) and speech comprehension (Peelle et al., 2013), suggesting that we may be more engaged with a stimulus when we synchronise to it.

Some research suggests that engagement and enjoyment may be related (Thompson, 2006; 2007). In support of this, Sachs et al. (2020) have also found that inter-subject synchronisation in the auditory domain, the default mode network, striatal circuitry, and orbitofrontal cortex correlate with enjoyment ratings when listening to music. However, Kaneshiro et al. (2020) did not find a relationship between ISC, engagement, and enjoyment.

Given this evidence, peripheral signals are a fitting way to assess the affective and engagement states of a listener objectively and continuously. However, the majority of these have been explored in the laboratory setting, with only a few studies that have moved to concert settings (Chabin et al., 2022;

Egermann et al., 2013). While there have been steps in understanding physiology as an implicit measure of emotion, to our knowledge there has been no study on engagement of music using physiological synchrony. With many studies looking at neural synchrony (Holroyd, 2022), especially for the phase domain, the question arises whether such methods can be insightful also for peripheral responses. As concerts permit a greater sense of absorption (Lamont, 2011), another natural open question emerges: how do these synchrony measures fit when audiences experience a more emotional and engaging setting? In the current thesis, these questions were explored further in a concert setting.

4 Overview

Within the field of cognitive neuroscience and neuroaesthetics of music, there is increasingly more evidence that the dynamic interplay of some musical features and the multimodal input in music performances can influence self-reported pleasure, liking, and aesthetic emotions as well as potential neural and physiological correlates of these responses (Brattico, 2021; Brattico et al., 2013; Brattico & Pearce, 2013; Cheung et al., 2019; Omigie et al., 2021). Although context plays a significant role in many empirical models of aesthetics, research beyond laboratory settings, extending to more ecologically valid settings, is still in its infancy (Brattico, 2021; Tervaniemi, 2023). To further the field's understanding of music listening in more real-world contexts, this thesis explores aesthetic responses – and their potential physiological correlates – in Western classical concerts. By focusing on two concert aspects – naturalistic music features and the multimodal aspects of a music performance – this thesis investigated their impact beyond laboratory settings while attempting to replicate laboratory results in concert settings. In particular, open questions of naturalistic music listening pertain to the length of music listening as well as the interplay of several features. A further overall aim of the thesis was to evaluate whether (synchrony of) physiological responses might complement investigating music listening in a concert setting. To this end, two datasets of concert series collected in 2018 were analysed and are presented in Chapters 2-5.

Chapters 2 and 3 present data analysed from string quintet concerts performed in October 2018 on consecutive days (2018.10.28, 2018.10.29, 2018.10.30), where both self-reports and physiological measures were collected. Although more data collection for concerts was planned, due to the COVID-19 pandemic, another previously collected data set was used for the experiments presented in Chapters 4 and 5. These were piano concerts: two concerts in March (2018.03.02, 2018.03.06) where only self-reports were collected and an additional two concerts in June (2018.06.11, 2018.06.12), where both self-reports and physiological responses were collected. These concerts included an experimental manipulation: participants heard piano performances either in an audio-visual (AV) condition or audio-only (AO) condition. Both of these concerts had a music program consisting of two common practice period (CPP)[8] pieces (i.e., Baroque, Classical and Romantic styles), with conventional musical structures, tonal harmony, and regular metres, and one contemporary classical piece with non-tonal harmony and irregular meter.

Chapter 2 assessed a more general aesthetic experience of the concert setting, in particular with regard to the global structure of the concert and full-length musical pieces, employing both self-reports and physiological responses as measures of an aesthetic experience. Chapter 3 investigated physiological responses over the time-course of music, using synchrony as a means to assess systematic responses to local musical features. Chapters 4 and 5 assessed aesthetic responses and physiological differences between modalities (AV and AO). Chapter 4 presents a more general approach in observing self-reported and physiological differences between modalities, while Chapter 5 explores differences between modalities in terms of real-time engagement, indexed by physiological synchrony. Chapter 6 provides a discussion of the empirical results (Chapters 2-5), evaluating what new insights can be gained by testing music experiences in such a setting, whether they could replicate laboratory results in a concert setting, and the extent that physiological responses could be used to add insight into such an experience. The limitations of such research are then presented, followed by proposals for potential directions for future research.

[8] Western classical music tradition based on the 12-tone tonal (major-minor) system, spanning from the mid Baroque to the late Romantic (approximately 1650- 1900).

Chapter 2

Aesthetic experiences of live concerts

Based on:

Merrill, J., **Czepiel, A.**, Fink, L. T., Toelle, J., & Wald-Fuhrmann, M. (2021). The aesthetic experience of live concerts: Self-reports and psychophysiology. *Psychology of Aesthetics, Creativity, and the Arts*, *17*(2), 134–151. https://doi.org/10.1037/aca0000390

Abstract

Music listening can lead to strong aesthetic experiences. However, to gain deeper insights into such experiences, more empirical research outside of laboratory settings is required. The current exploratory study measured aesthetic experience (music-induced emotions and absorption) in combination with psychophysiology (facial electromyography and arousal measures) from 98 participants during three live concerts with a program of classical, romantic, and contemporary chamber music. One musical movement from the contemporary work was presented from a recording. Results first highlight two key components of the concert frame as influencing the aesthetic experience. The programming order led to a rise-and-fall trajectory of emotions with the less familiar contemporary work leading to higher negatively valenced emotions. Nonetheless, this experience was embedded in an overall highly appreciated concert, with the factor of liveness becoming apparent in lower engagement with the recorded than the live music. Second, the participants' reactions gave insights into how the multimovement works were perceived; opening and closing movements elicited higher positively valenced arousal, contrasting the characteristics of an inner section, which evoked lower arousal and mixed emotions. This scheme differed between the classical and the romantic works in the third movements, reflecting a different trajectory of tension and relaxation in the respective styles. Finally, we show relations between physiological responses and self-reports reflecting both positive and negative aesthetic experiences. Overall, we demonstrate that the ecological validity of the current study is particularly informative for theoretical approaches to the aesthetic experience, with the frame as a crucial component.

1 Introduction

Music listening can lead to strong aesthetic experiences (Gabrielsson & Wik, 2003; Panzarella, 1980). While current frameworks of aesthetic experiences and emotions in music exist (Brattico et al., 2013; Brattico & Pearce, 2013; Hargreaves & North, 2013; Juslin, 2013), there is still a question of ecological validity in terms of stimuli and listening situations. So far, studies have typically been conducted in highly controlled laboratory environments, using only short musical works or excerpts (single movements or short clips). While such research may provide a good basis for understanding fundamental relationships between acoustic and musical characteristics, self-reports, and physiological responses, it is far from naturalistic forms and situations of music listening. In particular, typical lab approaches often do not include entire (multi-movement) pieces nor capture the supposedly crucial influence of the frame (Goffman, 1974) in which listening happens, despite situation and context being theoretically acknowledged as a key component in models of a visual aesthetic experience in art (e.g., Leder & Nadal, 2014; Locher et al., 2010; Pelowski et al., 2016) and in the reciprocal feedback model in music (Hargreaves et al., 2005). In real life, the frame codetermines (among others) the intended function of the music, the works that are selected in order to achieve this function, the adopted listening mode (e.g., attentive or not), or what other behaviours are deemed appropriate in response to the music (Greb et al., 2018; Schäfer & Sedlmeier, 2009). Therefore, the current study aimed to explore the aesthetic experience of music listening in a real-world setting, namely in response to live performances of music in a concert hall including three string quintets by Ludwig van Beethoven, Johannes Brahms, and the contemporary composer Brett Dean.

The Concert Frame: Attentive Listening in a Set Course

To describe the relationship between situational factors and music-related behaviours, we adopt a term introduced by the sociologist Erving Goffman, that is, the frame (Goffman, 1974). According to Goffman, frames contain components that are used by participants to understand, interpret, and evaluate the present situation and to align their behaviour accordingly. Therefore, these frame components can be expected to also influence aesthetic experience and appreciation. Recently, studies have started to acknowledge the influence of the context, situation, or frame by moving into real-world settings such as theatres (Ardizzi et al., 2020), movie theatres (Kostoulas et al., 2015), art museums (Pelowski et al., 2017; Tschacher et al., 2012), and live music venues (Balteş & Miu, 2014; Coutinho & Scherer, 2017b; Egermann et al., 2013; Egermann & Reuben, 2020; Scherer et al., 2019a)

A classical concert, which is the highly ritualised product of social conventions and stylistic development of Western music (Heister, 1984; Hewett, 2017), contains several relevant frame components. These include the programming (deliberately choosing popular/famous works from the classical canon and perhaps a less well-known piece; Gilmore, 1993; Weber, 2003) the multi-modal nature of the event (one cannot only hear the music live, but also watch the musicians creating it in the moment), and the setting aimed at affording an undisturbed, attentive, even absorbed mode of listening (created by the typical concert-hall layout with seats facing a lit stage, optimal acoustics, and learned behavioural regimes that call for silent and still listening; (Heister, 1984; Seibert et al., 2020; Small, 1998; Thorau & Ziemer, 2019). Accordingly, the concert frame signals to the audience that the works to be heard are of high artistic value, and that they therefore should be received with attention because they promise the listener a special experience.

Indeed, studies have shown that certain concert components enhance the music listening experience. Live performances of music tend to more frequently evoke strong (positive) responses (Gabrielsson & Lindström Wik, 2003; Lamont, 2011). Additionally, a live performance in a church venue was shown to induce higher emotional engagement, and feelings of wonder and tenderness, compared to a video recording of the same performance shown in a lecture hall (Coutinho & Scherer, 2017b). Attention has also been shown to be related to more intense, even absorbed experiences (Lange et al., 2017; Vroegh, 2018). Also, when asked about their motivation to go to concerts, attendees typically refer to their hope for deep and meaningful experiences (Brown & Knox, 2017; Burland & Pitts, 2014; Holt, 2010). Although studies have begun showing the effect of attentive listening during live opera performances (Balteș & Miu, 2014; Scherer et al., 2019a), no studies exist that explore the trajectory of an aesthetic experience with an intentional series of differing multi-movement works, which are a key component of classical concerts.

The Compositional Strategy: A Music Theoretical Perspective

While the concert frame determines the schematic nature of the concert on a global level, there is also the possibility to consider smaller units within the single compositions. To investigate these underlying compositional strategies in music, a music theoretical perspective can be taken on (Zbikowski, 2002). In a similar way to how linguistic research demonstrates that lower-level units group to higher-level structures – syllables group to words (with attentive listening; Ding et al., 2018), and words group to hierarchical sentence structures (Nelson et al., 2017) – such grouping is present also in music (for a review, see Patel, 2003), and systematic music theoretical methods further suggest categories to describe longer-term musical features. These longer-term features – such as themes and phrases that create sections – eventually constitute movements (i.e., separate parts in multimovement works such as sonatas or symphonies). These categories are not only mere descriptions of the musical score, but have assigned functions that they fulfil for the listener within a given style (Caplin, 1998). Like the concert as such, compositions also follow certain trajectories that are typical for their respective style and genre.

Those familiar with European music between the 18th and the early 20th centuries (the so-called common practice period) in general have heard a comparable array of music – which follows a similar compositional strategy – many times, to the extent that an expectation of how a (full-length) work and its movements may sound is developed. This is the basic assumption of a theory on one of the most influential and discussed forms in classical music: the sonata form (Hepokoski & Darcy, 2006). Themes and movements are presented in a predefined order and harmonic schema, following specific functions by raising or denying expectations to create a higher-level form. This convention is typically used in symphonies and sonatas as well as chamber music genres of the (Viennese) classical and romantic period, which are represented in the current study through Ludwig van Beethoven's and Johannes Brahms' works, respectively. The traditional sonata cycle follows a standard four-movement pattern, with different tempi for each movement (fast–slow–dancelike–fast) as well as a development of key areas and the resulting harmonic expectancy (tonic–nontonic–tonic–tonic). Fast opening and closing movements frame two inner movements of a contrasting slow, song-like character, and a rhythmically pronounced dance-like character in triple meter (a minuet or scherzo). The aggregation of these movements has recently been described as the pairwise grouping of two movements toward a balanced symmetry of 2 + 2 (Hepokoski & Darcy, 2006, p. 338). However, empirical evidence whether this grouping is actually perceived by listeners or only part of a theoretical layout is still pending.

Although listeners implicitly gain knowledge through exposure (Bigand & Poulin-Charronnat, 2006) and have implicit knowledge of tonal music (Koelsch et al., 2000; Tramo et al., 2001) – at least on a local phrase level – psychological research does not seem to provide perceptual evidence for such high-level features of musical form. Concerning whole sections, studies have shown that aesthetic pleasure does not change when the intended order of form parts is not always in place, that is, comparing stimuli in original versus manipulated forms (Cook, 1987; Gotlieb & Konečni, 1985; Karno & Konečni, 1992). However, emotional characteristics and physiology related to the formal functions of whole movements from different musical styles has not been investigated.

In our concerts, two instantiations of a sonata cycle (Beethoven and Brahms, both four movements) were contrasted by the four sequential movements of a "suite" (Brett Dean, contemporary work). This allowed for investigating the audience's responses to the formal structure of two different four-movement patterns: the cyclic organisation of the sonata form and the sequential order of the suite.

Measuring the Aesthetic Experience

Although there are several dimensions that can encompass an aesthetic experience, we chose to focus on two components that have been previously identified as key components of an aesthetic experience with music (Brattico et al., 2013; Brattico & Pearce, 2013), that is, music-evoked emotions and absorption.

The importance of absorption in the aesthetic discourse about classical music has been discussed in music history literature (Herbert, 2016; Høffding, 2018). Conceptualising the aesthetic experience as a distinctive state of mind (Levinson, 2003; p. 7; Vroegh, 2018) that is related to the focused perception of sensory objects and differs from "everyday" mental states, the state of absorption presents itself as a valuable measure for the aesthetic experience of music in a concert, where parameters are traditionally designed to optimise focus only on the music. These parameters include dimmed lights, an engineered room architecture for optimum acoustics (which influence perceived clarity of melodic lines, loudness, and reverberation; Lokki et al., 2011, 2012) as well as behavioural conventions of attentive listening such as (expected) total silence and sitting still (Heister, 1983).

Emotions have also been shown to be part of an aesthetic experience (Brattico et al., 2013; Brattico & Pearce, 2013; Menninghaus et al., 2019; Schindler et al., 2017). It is worth noting here the differentiation between perceived and felt emotions (Gabrielsson, 2001; Schubert, 2013). While works of classical music are expected to have an expressive component and to be performed expressively, they do not necessarily evoke the same emotions that they might express, nor are they expected to (e.g., music may evoke mixed emotions such as being moved; Trost et al., 2012). Recent research has led to more standardised self-report tools to measure such evoked emotions, with some constructed to be applicable to art and nonart (e.g., nature) domains, such as AESTHEMOS (Schindler et al., 2017), or shorter tools that apply to more specific domains such as classical music (GEMIAC; Coutinho & Scherer, 2017a).

In addition to self-reports, a nonobtrusive and continuous method to measure experienced emotions is using electromyography (EMG), cardiovascular and respiratory measures, and skin conductance (see Cacioppo et al., 2000). Responses such as increased sweat gland secretion (measured by skin conductance), heart and respiration rate – activated by the sympathetic division of the autonomic nervous system in response to stimulating events – reflect increased felt arousal and engagement (Benedek & Kaernbach, 2010; Bradley & Lang, 2000; Cacioppo et al., 2000). Facial muscle activity is also considered a measure of behaviour (Cacioppo et al., 2000) and can reflect mimicry of several discrete emotions (Wingenbach et al., 2020) and – to a certain extent – evoked emotions, for example, the zygomaticus major (smiling) muscle and corrugator supercilii (frowning) muscle have been shown to reflect positive and negative valence, respectively (Bradley & Lang, 2000).

Most research on physiology and emotions during music listening concerns how the perceived emotion of the music is reflected in felt emotions, that is, the emotion of the music is induced within the listener. Studies using either categorical (e.g., happy, sad) or dimensional (valence-arousal) emotion concepts show that “happy” or arousing music (rated as inducing happiness or high arousal in listeners) evoked higher skin conductance amplitudes and faster breathing and heart rates compared to sad music or music with low arousal (happy vs. sad: Etzel et al., 2006; Khalfa et al., 2002, 2008; Krumhansl, 1997; Lundqvist et al., 2008; high vs. low arousal: (Egermann et al., 2013; Gomez & Danuser, 2004). Higher smiling muscle activity was also associated with happy compared to sad music (Khalfa et al., 2008; Lundqvist et al., 2008).

Other research demonstrates how physiology may reflect emotions above and beyond the emotion of the music itself, such as pleasure, aesthetic chills, or perceived beauty. Music rated as pleasurable, or reported as evoking chills, evoked higher skin conductance amplitudes and faster breathing and heart rates (pleasure: Salimpoor et al., 2009; Sammler et al., 2007, chills: Blood & Zatorre, 2001; Craig, 2005; Guhn et al., 2007; Rickard, 2004; Roy et al., 2009; Salimpoor et al., 2009). Pleasant music also reduces certain motor activity such as eyeblink amplitude (Roy et al., 2009). Additionally, heart rate is found to be higher during music judged as “beautiful” (although this effect is dependent on speed; de Jong et al., 1973). More recently, Omigie et al. (2021) found an interaction between the music judged as beautiful and its induced emotions: higher skin conductance was found in high energy beautiful music, whereas more smiling and higher respiration rate were associated with low tension and low energy beautiful music. However, there is a limit to studies that show which kind of physiology may reflect more negative aesthetic experiences. Particularly in a live music situation, where full pieces of music are presented, one can expect a variety of music-evoked emotions.

The Present Investigation

This current study considers specific concert features and full- length musical trajectories of differing genres as important elements to influence effects on aesthetic experiences (here measured using behavioural responses and psychophysiological measures), thus making our approaches and hypotheses rather exploratory. To investigate the aesthetic experience of music in a real-world (as opposed to laboratory) situation, the current study tested 98 participants across three identical live concert evenings. The concert program comprised three string quintets in classical, romantic, and contemporary style with four movements each. Participants were asked to fill out short questionnaires on their aesthetic experiences, that is, music-induced emotions (GEMIAC; Coutinho & Scherer, 2017a) and absorption (Vroegh, 2018), which were reported after each movement. Physiology was continuously measured throughout the whole concert.

Despite the exploratory nature of this ecological set-up, we nonetheless had a few hypotheses. We expected to (1) see how behaviour and physiology may reflect the concert frame, for example, its programming in the form of differences between the classical, contemporary, and romantic works. We also expected to (2) observe compositional strategies of the works mirroring in audience responses, that is, differences between movements due to their functions within a sonata cycle or suite. By choosing full pieces of music that were able to elicit an aesthetic experience on various levels – rather than choosing music that reflects one specific emotion (e.g., Khalfa et al., 2002, 2008) or that evokes primarily pleasurable emotions (e.g., Omigie et al., 2021; Salimpoor et al., 2009) – we additionally hoped to (3) extend typical laboratory studies on the relations between physiology and a wider range of positive and negative aesthetic experiences. This study will widen our knowledge on how people experience music aesthetically and emotionally based on self-reports and psychophysiology in the real-world setting of a live concert.

2 Method

Venue

The ArtLab is a special concert hall designed for empirical investigations. It comprises 46 comfortable seats with armrests and a pull-out table, arranged in five rows. The room is engineered to ensure equal sound quality for all seats. It has a stage, sound system with speakers, microphones and video cameras on the ceiling. In front of the ArtLab is a large foyer, where participants were prepared for physiological data acquisition.

Participants

Forty-six different participants responded to a public advertisement for a chamber music concert for each of the three concerts. Some participant data were lost due to technical issues of server and user failures (N = 31). A resulting number of 98 physiology data sets (Concert 1 (C1), N = 36; Concert 2 (C2), N = 41; Concert 3 (C3), N = 21) and 88 complete data sets (physiology and self- report data) were used for final analyses.

The gender distribution was similar across concerts (C1 = 15/17 female/male, C2 = 16/17, C3 = 9/12; 3 not stated). Age groups across concerts were also quite similarly distributed: Participants were asked to choose their age group (within a five-year range, i.e., 18–22, 23–27, . . . 95–99), where half of the participants in C1 marked their age as being between 18 and 50, in C2 between 18 and 55, and C3 between 18 and 40 years old. Most participants in all three concerts indicated their highest education level was a university degree (N = 23, 25, 17), a high school degree with German “Abitur” (N = 8, 5, 4), or a finished vocational/professional training (without “Abitur”; N = 1, 4, 0).

To evaluate musical sophistication, two subscales from the German version of The Goldsmiths Musical Sophistication Index (Schaal et al., 2014) were chosen: General Music Sophistication and Emotions (Müllensiefen et al., 2014). The participants across the three concerts were very similar a) for the overall sophistication: C1 (mean (M) = 69.84, SD = 22.01), C2 (M = 71.61, SD = 21.97), C3 (M = 70.76, SD = 19.33), and b) for the emotion scale: C1 (M=33.53, SD=4.75), C2 (M=31.17, SD=6.85), C3 (M= 33.24, SD = 5.45).

Participants reported that they most often visit concerts comprising classical music and opera. Half of the participants came to the concert with someone else (see online supplemental materials for details).

Materials

The chosen repertoire for the live performances of a string quintet was based on an artistic decision by musicologists and curators, reflecting the stylistic variety and order (including an interval) typical of present-day classical chamber music concert programming. The first work played was from the Classical period, op. 104 in C minor (1817) by Ludwig van Beethoven, with four movements. From the second work, “Epitaphs” (2010) by the contemporary composer Brett Dean, also four movements were presented – the fifth one was omitted for artistic reasons. The third work was from the Romantic period: the string quintet op. 111 in G major (1890) by Johannes Brahms, also comprising four movements.

The contemporary work by Dean is written as a suite with five movements, each of them musically independent from each other, and each paying homage to personal friends. This work generally has more complex/dissonant harmonic and rhythmic structures. Each movement emphasises a characteristic set of acoustic features (see Figure 1), some of them more challenging to the standard listening experience than others. For instance, timbre and intonation of the performance of the first movement – resulting from the realisation of various harmonics in the score – can be described as more challenging to listen to than the more consonant and melodic third movement. The last (fourth) movement played in the concert has the character of a typical finale in terms of clear, engaging rhythm, and repetitive structures. Even though there is no standard musical form to label the overall work, there is a clear dramaturgical development from movement one to movement four, which closes with an effectual ending. Beethoven's and Brahms' quintets are exemplars of the traditional sonata cycle (Hepokoski & Darcy, 2006), following the standard four-movement pattern as described above (see Figure 1).

The performing musicians were a string quintet from the Frankfurt am Main (Germany) area, working solely as professional orchestra and chamber musicians. The musicians were accordingly compensated for their time playing in all three concerts. As the main aim was to keep the three performances as comparable as possible, the musicians were instructed to achieve similarity across concerts, which was more likely to be ensured with a professional group of musicians that regularly play together.

The second movement of the work by Dean, "II. Walk a Little Way With Me" was presented as a recording (Doric String Quartet, 2015; Brett Dean, viola) via loudspeaker (with musicians still sitting on stage). The idea behind this decision was to have one stimulus presentation with the exact same acoustic properties across all three concerts and without the visual aspects and the live acoustics, which are (besides the frame) typical characteristics of laboratory situations. As one of the main aims of the study was to keep the setting as naturalistic as possible, it seemed most plausible for the concert to have the control comparison recording in the contemporary work – as this style typically can also include electro-acoustic elements – rather than in the classical work or the romantic work (which may lead to confusion if a recording of only one of the movements was played throughout a classical or romantic work).

Work		Ludwig van Beethoven String Quintet in C minor op. 104	Brett Dean 'Epitaphs'	Johannes Brahms String Quintet No. 2 in G major, op. 111
Stylistic period, date of composition		Classical (1817)	Contemporary (2010)	Romantic (1890)
Overall musical form		Sonata form	Suite	Sonata form
Movements				
Title or tempo description Musical characteristics (tonality and key area, meter, timbre, etc.)	1	Allegro con brio C minor – tonic 3/4	'Only I will know'. Gently flowing, with intimate intensity. 4/4 infused by 5/8 and 6/8. Unique texture through the use of overtones and bowing techniques.	Allegro non troppo, ma con brio G major – tonic 9/8
	2	Andante cantabile con variazioni E-flat major – non-tonic 2/4	'Walk a little way with me'. Moderato scorrevole. * Subsequent sections with varying meter and tempo.	Adagio D minor – non-tonic 2/4
	3	Menuetto Quasi allegro C minor (Trio in C major) - tonic 3/4	'The philosopher'. Quasi cadenza – suddenly flowing and floating. No meter in the beginning, then steady pulse, no clear pulse in the end.	Un poco allegretto G minor (Trio in G major) – non-tonic 3/4
	4	Finale. Prestissimo C minor – tonic 2/2	'György meets the Girl Photographer'. Fresh, energetic. Steady pulse of 1/16 notes with changing accents, engaging rhythm.	Vivace ma non troppo presto – Animato G major – tonic 2/4
	5		(not presented in the concert)	

Figure 1. Program and Description of The Works

Note. The term "work" describes the specific performance of that piece as experienced by the participants of the current study. * Presented via loudspeakers.

Procedure

Participants were invited to arrive at the venue one or one and a half hour(s) prior to the beginning of the concert. First, participants were informed about the study and signed required agreements to take part in the study. They were also handed a program that included the composer's names, the works with year of composition, and the names of the movements (no further text was presented). They were then prepared for the physiological recording, which included the measurement of the (a) zygomaticus major (smiling) muscle activity measured by EMG with adhesive electrodes on the left side of the face (with a ground electrode placed on the mastoid), (b) respiration rate (RR), measured using a custom prepared respiration belt wrapped around the lower rib cage, (c) electrodermal activity (EDA), measured with electrodes attached to the index and middle fingers of the nondominant hand – from which phasic skin conductance (SC) could be inferred – as well as (d) blood volume pulse measured with a plethysmograph clip – from which heart rate (HR) could be inferred. Electrodes were connected to a portable amplifier, the "biosignalsplux" system (https://plux.info/12-biosignalsplux; "plux" hereafter). The plux device recorded the data from the incoming sensors and was attached with adhesive tape to the participant's upper back and remained there for the entirety of the concert. Data were sampled at a 1000 Hz rate. Concerts started at 7:30 p.m. Beethoven and Dean were performed in the first half; Brahms was performed in the second half. In between, a 20-minute break was taken outside the ArtLab concert hall foyer. Concerts ended approximately 9:45 p.m.

After each movement, a short break (roughly two minutes) was taken to fill out two short questionnaires on a tablet. The first questionnaire comprised eight items of state absorption taken from Vroegh (2018, p. 150 and p. 203) including liking, all evaluated on a 5-point Likert scale from very much to not at all. The questionnaire comprised items on altered awareness, such as "I was completely absorbed by the music," "Time passed quickly"; on dissociation, such as "I did not notice the surroundings," "I forgot being at a concert"; attention, such as "I focused completely on the music"; and control items, such as "My mind was wandering" and "I was totally bored." The second questionnaire on the next page comprised the 14 "emotion classes" from the GEMIAC (Coutinho & Scherer, 2017a). This "checklist" contains a selection of affect and emotion categories, where each category was derived from a hierarchical cluster analysis, leading to so-called emotion classes. Each class is represented by one item containing two terms: filled with wonder/amazed, moved/touched, enchanted/in awe, inspired/enthusiastic, energetic/lively, joyful/wanting to dance, powerful/strong, full of tenderness/warmhearted, relaxed/peaceful, melancholic/sad, nostalgic/sentimental, indifferent/bored, tense/uneasy, agitated/aggressive. The items were translated into German by taking the adjectives from a comparable assessment, AESTHEMOS (Schindler et al., 2017), and the intensity of each emotion as felt by participants were assessed on a 5-point Likert scale.

Additional questions were presented after each work and after the concert, for example on familiarity with this kind of music (options of "Yes," "No," and "I am not sure"; see online supplemental materials). Participants were compensated with a free ticket to the concert, including a complementary nonalcoholic/-caffeine refreshment during the interval. They did not receive any further monetary compensation. The three concerts were conducted on three consecutive evenings. In order to have the concerts as comparable as possible, great care was taken to ensure factors of timing, performers, lighting, and temperature, which were kept the same across these three concerts.

Analysis of Physiological Measures

All signal processing was carried out using custom scripts written in MATLAB 2018b (The MathWorks). Plux data were transformed from raw to meaningful measures following the data sheets provided by the developer, that is, EMG and BVP in millivolt, EDA in microsiemens, and respiration in percentage of displacement. Missing data (gaps of less than 50 ms) were interpolated at the original sample rate.

Electromyography

To obtain a continuous measure of facial muscle activity over time, the EMG was band-pass filtered between 90 and 130 Hz, and the absolute value of the Hilbert transform of the filtered signal was extracted and then smoothed using the conv2 function in MATLAB, as recommended (http://www.fieldtriptoolbox.org/ documentation; Oostenveld et al., 2011).

Respiration and Heart Rate

The respiration signal was low pass filtered by 0.6 Hz and demeaned. BVP signal was band-pass filtered between 0.8 and 20 Hz, and demeaned. For both signals, peaks in the raw signal were identified using a function that identified local minima and maxima in any given epoch, and visual inspection was used to confirm that peaks were correctly identified. Finally, rate for each signal was estimated by taking a differential of the timings of maxima (interbeat intervals), and the resulting time-series were interpolated to a regular sampling frequency to give continuous HR and RR. Data were kept at 1000 Hz sampling rate. Noisy HR data were additionally identified using Poincaré plots (Mourot et al., 2004; Sosnowski et al., 2005), and single work movements by single participants were excluded if more than 10% of data were affected.

Skin Conductance

EDA was down sampled to 20 Hz and analysed by means of Discrete Decomposition Analysis from the Ledalab toolbox (Benedek & Kaernbach, 2010) performing optimization of four initial values and data smoothing using the Gauss method and a window width of 16 samples. This resulted in a mean amplitude of phasic skin conductance components (SC), which reacts in an event-related manner.

Measures

In order to get meaningful measures despite varying movement lengths, number of events per minute was used, in addition to mean rate and amplitude. For SC, the mean amplitude of phasic activity and number of events per minute were calculated; however, the number of events did not follow a normal distribution, so only the mean amplitude of phasic activity was taken for further analysis. For the EMG signal, mean amplitude did not give a meaningful measure, as the amplitude averaged out too much (most likely as the time across each movement was several minutes, as opposed to typical trials lasting a few seconds), so the number of events per minute was used instead. An event was defined by moments when the signal was higher than a threshold (the individual mean plus two times the standard deviation was used, i.e., being the threshold for a $p < .05$ value following a normal Gaussian distribution). For HR and RR, the mean per movement was calculated. This yielded four physiological data points per participant per movement: mean HR (beats per minute, BPM), mean RR (breaths per minute), mean amplitude of phasic skin conductance, and mean number of EMG events per minute, although after manual data checks some participants' HR and/or SC data were rejected, so some participants only had two or three physiological data points. In total, 98 data sets were used for EMG and RR, 88 for skin conductance, and 82 for HR data.

Acoustical Analysis

To compare acoustic variability between the concerts, acoustic features representing performance-based features (i.e., features not describing the music composition itself, but rather acoustic features that may vary from performance to performance; Beveridge & Knox, 2009; Goodchild et al., 2019) were obtained manually and computationally. Using the MIRToolbox (Lartillot & Toiviainen, 2007), RMS (Root Mean Square energy, a feature typically representing loudness perception) and spectral centroid (the feature most commonly used to represent timbre, or "brightness"; Alluri et al., 2012; Grey & Gordon, 1978; Iverson & Krumhansl, 1993) were extracted in 25-ms time windows as is typical in MIR analysis (Tzanetakis & Cook, 2002). As a longer-term feature, tempo is not always successfully captured by computational methods (Lange & Frieler, 2018) and was therefore extracted manually using Sonic Visualiser (Cannam et al., 2010), by tapping in timestamps for each beat of the performances of each movement (yielding a vector of timestamps per beat and bar), calculating the interonset intervals between beats, and converting these to beats per minute. Loudness, timbre, and tempo values were averaged in time windows per bar to allow comparisons between concerts.

Statistical Analysis

Statistical analyses were carried out using R (R Core Team, 2014). All questionnaire items were recoded so that the highest number reflects highest intensity. The *lme4* package within R (Bates et al., 2015) was used to construct linear mixed effects models (LMMs). The significance of the fixed effects was tested with conditional F-tests using the *anova* function and the *lmerTest* package (Kuznetsova et al., 2017). Pairwise comparisons using the Tukey method were performed with the emmeans package (Lenth, 2021). Two (Pseudo-)R^2 values (marginal and conditional) were calculated with the MuMin package (Barton, 2015), which is designed for general and linear mixed effects models. The randomForest package in R (Liaw & Wiener, 2002) was used to construct random forest regressions. Five hundred trees were fitted. The output gives the percent explained variance of each model and the importance of each variable in explaining the variance. The "importance" of each predictor was indicated in %IncMSE, which shows the increase in mean squared error of predictions as a result of a variable being permuted. Because random estimations depend upon the computer's random number generator, each model was replicated 100 times to examine the distribution of R^2_{oob} as well as the variable importance across replications. For multiple and forest regressions, reference values for interpreting R^2_{GLMM} and R^2_{oob}, respectively, can be taken from Cohen (1988), that is, a small effect $R^2 = .02$, medium effect $R^2 = .13$, and large effect $R^2 = .26$. All statistical results of fixed and random effects can be found in the online supplemental materials.

Acoustic Comparisons

Performance acoustic features were correlated per movement between concerts using the corr.test function from the psych package in R (Revelle, 2021), adjusted for false discovery rate. All acoustic correlations of instantaneous tempo, timbre, and loudness between concert 1 (C1), concert 2 (C2), and concert 3 (C3; C1–C2, C1–C3, C2–C3) had r values above .6, with $p < .001$ (with most correlations at $r > .8$), meeting the criteria for a large effect size in correlation of $r > .5$ (Cohen's q for correlation coefficients; Cohen, 1988). This confirms that the concerts were similar enough acoustically to allow further statistical analysis (i.e., that measured self-report and physiology across concerts were in response to a stimulus that remained comparable over the concerts). Unsurprisingly, apart from the (recorded) Dean 2nd movement, *r* values were under .95, suggesting that the performances differ slightly from each other, reflecting the "live" aspect of performances: that even highly practised performances offer slight deviations in interpretation to offer a fresher performance as well as exhibiting small human errors (Chaffin et al., 2007; Palmer, 1997).

Reliability of the Questionnaires

Intraclass correlation (ICC; two-way random effects model, type: consistency) was performed for the eight absorption items and revealed an average measure ICC of .535 with a 95% confidence interval from .491 to .577 ($F(1020, 7140) = 2.151$, $p < .001$). The ICC for the 14 emotion items from the GEMIAC revealed an average measure ICC of .734 with a 95% confidence interval from .709 to .759 ($F(924, 12012) = 3.766$, $p < .001$).

Factor Analysis

All items from the GEMIAC and the absorption scale were subjected to a factor analysis. The *fa* function (also from the psych package) was used with direct oblimin rotation, after determining the number of factors performing a parallel analysis. Five factors (latent variables) were identified (see Table 1).

Effects of Work and Movement on Emotion and Absorption Items

To investigate effects of performances of different works and movements on the self-reports, LMMs were performed for each latent variable separately as the dependent variable, with the fixed effect of (a) Work (three works, with random intercept of Movement and Participant) or (b) Movement (12 movements, with random intercept of Participant).

Effects of Concert, Work, and Movement on Psychophysiology

To evaluate effects of concert, musical work, and movement on physiological measures, separate LMMs were fitted for each physiological measure with fixed effect of Concert (three concerts) or Work (three works with random intercept of Movement and Participant) or Movement (12 movements, with random intercept of Participant) separately.

Effects of Emotion and Absorption Items on Psychophysiology

Two approaches were chosen to investigate the relationships between ratings and physiological measures. First, an exploratory approach using random forest regression was used to find the best predictors (based on importance values) within the self-reports for each psychophysiological measure. Random forests were fitted for each physiological measure with all rating items (N = 22) as predictors. Second – because the results from random forest regressions do not give an objective threshold for importance values – LMMs were fitted for each physiological measure with the latent variables as fixed effects (four physiological measures, with random intercept of Participant). Because one absorption factor and one emotion factor correlated (factors positive emotions and dissociation; $r = .57$), we made two models to avoid issues with multicollinearity. One model entailed the emotion-related factors (positive, negative, and mixed emotions as fixed effects), and the other the absorption-related factors (engagement and dissociation as fixed effects).

Table 1. Factor Structure of the Items from the GEMIAC and the Absorption Scale

Questionnaire	Items	Positive emotions	Engagement	Negative emotions	Mixed emotions	Dissociation
	Proportion of variance explained	0.29	0.22	0.19	0.18	0.12
GEMIAC	energetic/lively	0.82				
GEMIAC	powerful/strong	0.69				
GEMIAC	inspired/enthusiastic	0.69				
GEMIAC	joyful/wanting to dance	0.66				
GEMIAC	filled with wonder/amazed	0.62				
GEMIAC	enchanted/in awe	0.51			0.36	
Absorption	bored		−0.68			
GEMIAC	indifferent/bored		−0.67			
Absorption	concentrated		0.58			
Absorption	forgetting time		0.55			
Absorption	absorbed		0.54			0.37
Absorption	liking		0.47	−0.40		
Absorption	mind wandering		−0.31			
GEMIAC	tense/uneasy			0.89		
GEMIAC	agitated/aggressive			0.84		
GEMIAC	relaxed/peaceful			−0.43	0.41	
GEMIAC	nostalgic/sentimental				0.78	
GEMIAC	melancholic/sad				0.72	
GEMIAC	tenderness/warmhearted	0.30			0.51	
GEMIAC	moved/touched	0.42			0.46	
Absorption	forgetting being in a concert					0.73
Absorption	forgetting surroundings					0.70

Note. Factor loadings < .3 are omitted.

3 Results

Self-Reports

Concert

Across all three concerts, participants liked the concert as a whole (M = 4.10, SD = .69) as well as the atmosphere (M = 3.82, SD = .83). They had the feeling they were actually visiting a concert (rather than being part of a study; M = 3.81, SD = 1.01), reporting low disturbance of the electrodes (M = 2.17, SD = 1.0). Reported duration of preparation of physiological set-up before the concert (M = 3.98, SD = .81), and of the concert itself (M = 4.22, SD = .70) was appropriate, and participants could imagine visiting such a research concert again (M = 4.64, SD = .64).

Works

The quintets by Brahms and Beethoven were familiar for the majority of participants (only 19% and 29% respectively reported they were not sure if they knew it), which was not the case for Dean’s work (98% reported not to know it). Liking ratings of the (whole) works were for Beethoven M = 4.09 (SD = .88), for Brahms M = 3.99 (SD = .78), and Dean M = 3.25 (SD = 1.18).

Differences Between Works and Movements in Self- Reports

Effects of Work and Movement on each latent variable can be seen in the significant F-tests resulting from the linear mixed models in Table 2.

Works. Compared to Brahms and Beethoven, Dean's work received lowest ratings of positive emotions (marginally missing significance to Beethoven) and mixed emotions (significantly lower than Brahms), but highest ratings of negative emotions (significantly higher than both other works). Brahms' quintet was rated highest on dissociation (significantly higher than Dean). No significant differences were found between works for engagement (pairwise comparisons are depicted in Figure 2; all statistics can be found in the online supplemental materials).

Movements. In positive emotions, a difference occurred between the second and the outer movements of the classical and the romantic works: Beethoven's 3rd and 4th movements had significantly higher ratings compared to the 2nd movement. Brahms' 4th movement had significantly higher ratings compared to the 1st through 3rd movements (see Figure 2). For Dean, the 4th movement had significantly higher ratings compared to the 1st and the 2nd in positive emotions. In mixed emotions, a reversed pattern could be seen where Beethoven's and Brahms' 2nd movements had significantly higher ratings compared to the 1st, 3rd, and 4th movements. For Dean, the 1st movement had significantly higher ratings compared to the 4th movement. In negative emotions, Dean's 1st movement was significantly higher than the 4th. No significant differences within the Beethoven and Brahms movements were found. In engagement, Beethoven's 3rd movement was significantly higher than the 2nd, and Brahms' 4th movement higher than the 1st, 2nd, and 3rd, being the highest in the whole concert. Dean's 4th movement was significantly higher than the 2nd. There were no significant differences between movements within works for dissociation. The part presented via loudspeaker, Dean's 2nd movement, received lowest ratings in engagement, which was significantly lower compared to Dean's 4th, Beethoven's 3rd and 4th, and Brahms' 4th movements.

Table 2. Results of the Conditional F-Tests Following the Linear Mixed Effects Models for Each Latent Variable and the Fixed Effects of Work or Movement

	Work				Movement			
Latent variable	$F(2, 9)$	p	$R^2_{(m)}$	$R^2_{(c)}$	$F(11, 815)$	p	$R^2_{(m)}$	$R^2_{(c)}$
Positive emotions	3.781	.064	.058	.454	19.261	<.001	.131	.447
Mixed emotions	6.075	.021	.153	.589	53.013	<.001	.275	.576
Negative emotions	71.251	<.001	.339	.552	65.421	<.001	.360	.551
Engagement	0.867	.452	.007	.340	5.333	<.001	.043	.338
Dissociation	6.702	.016	.013	.477	3.310	<.001	.021	.479

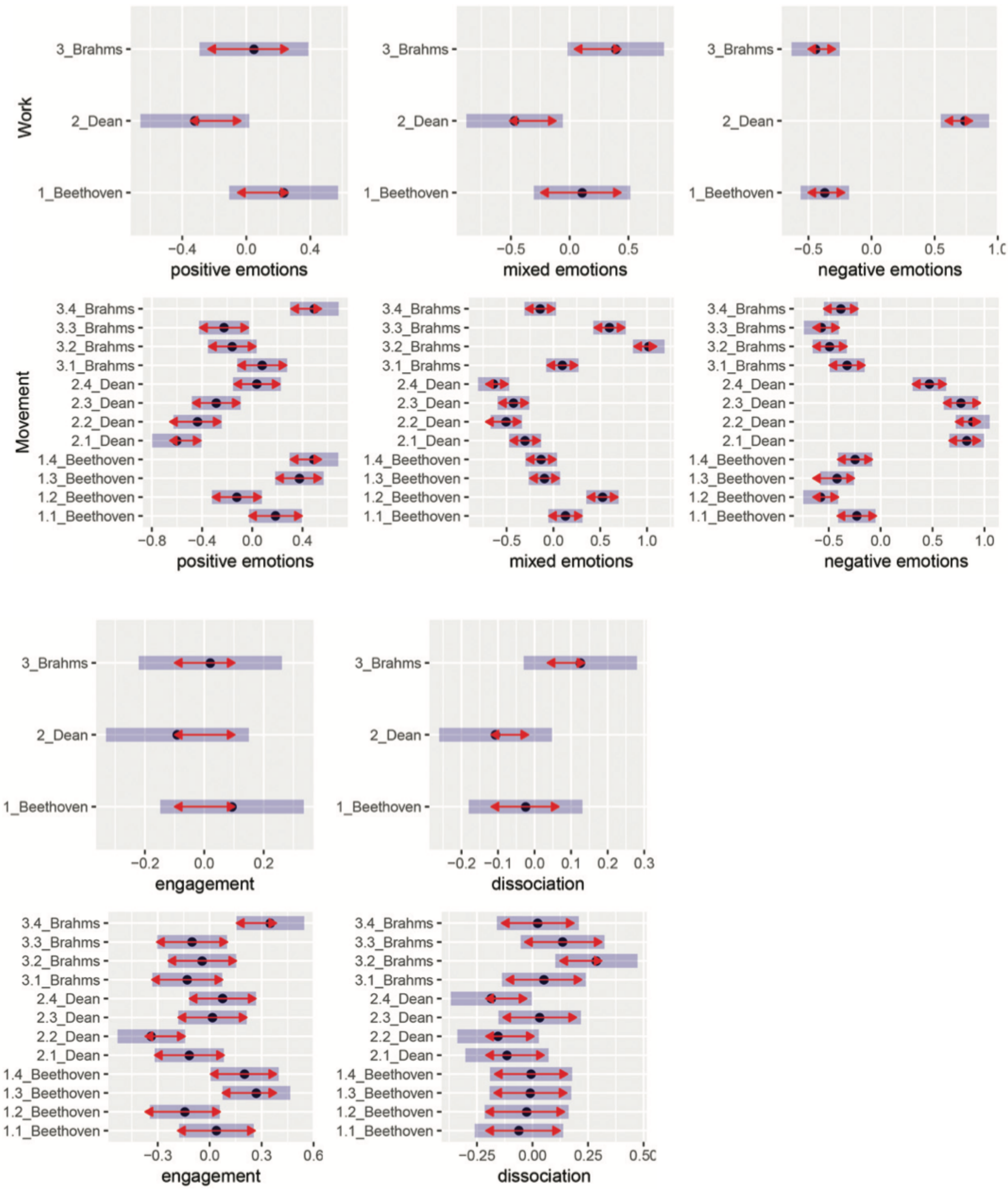


Figure 2. Differences (Estimated Marginal Means) Between Works (Upper Row) and Movements (Lower Row) in Emotion and Absorption Ratings (Latent Variables). Note. Pairwise comparisons with mean (black circle), 95% confidence intervals (CIs; blue (light grey) bars), and "emmeans" statistics depicting the significance of comparisons (overlapping red (dark grey) arrows mean no differences). See the online article for the colour version of this figure.

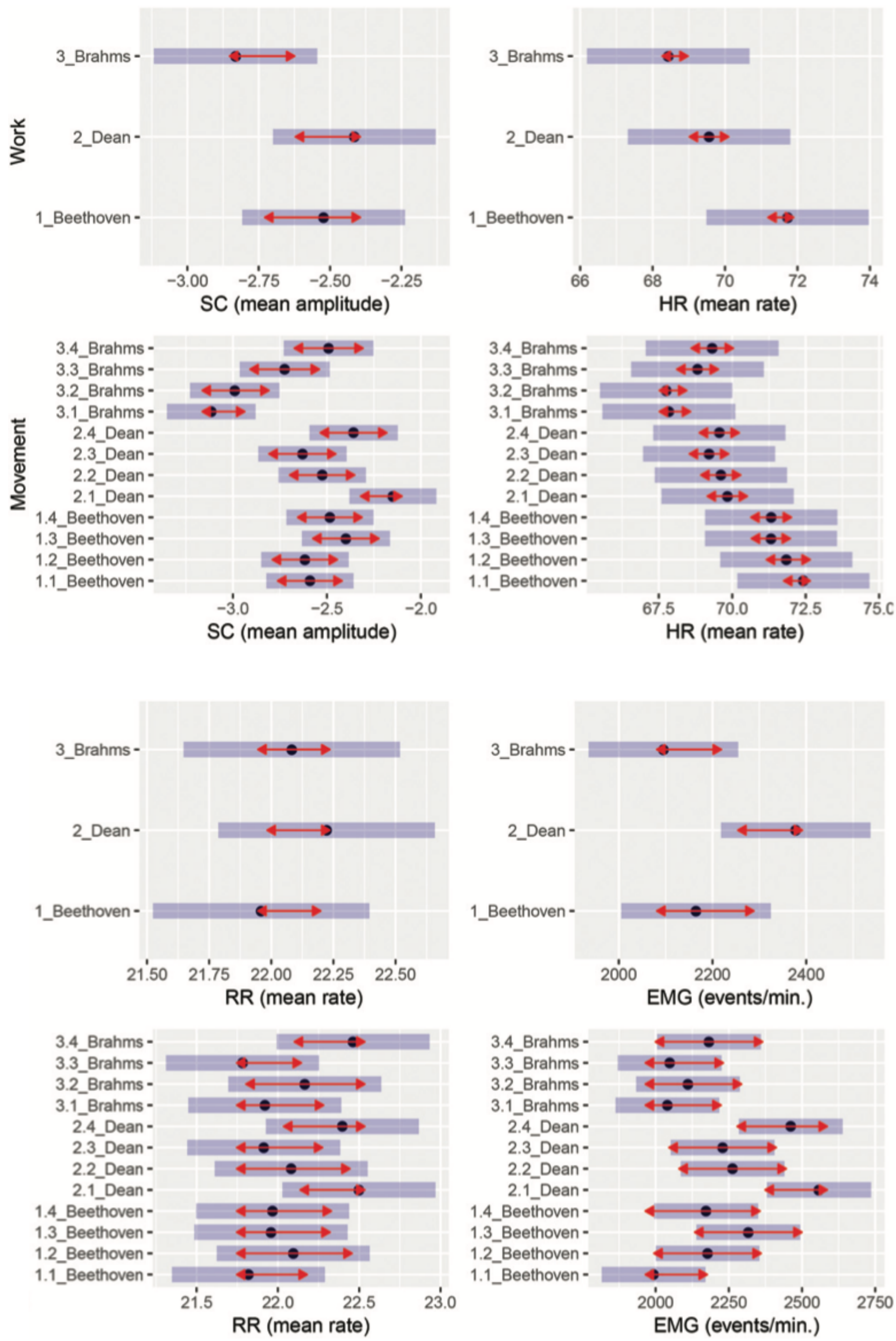


Figure 3. Differences Between Works (Top Row) and Movements (Bottom Row) in Psychophysiological Measures. Note. For details, see Figure 2. See the online article for the colour version of this figure.

Psychophysiological Measures

Differences Between Concerts, Works, and Movements

Concert. LMMs with physiological measures as dependent variables and concert as independent variable revealed no differences between concerts (see online supplemental materials). Along with the very similar acoustic features across concerts, this confirms that the physiological data from three separate, but comparable concerts do not differ significantly based on the mean values and can be treated as a single group. Therefore, no comparisons are made between concerts in the following analyses.

Works. The LMMs for Work show significant differences for SC, EMG, and HR response, but not for RR (Figure 3; Table 3). Dean received significantly higher physiological responses than Brahms for SC, HR, and EMG. Beethoven had significantly higher HR responses compared to Dean and Brahms.

Movements. The LMMs for Movement show significant effects for SC, EMG, and HR response, but not for RR (see Figure 3). SC is significantly higher in Dean's 1st movement than in Dean's 2nd and 3rd movements, Brahms' 1st through 3rd movements, and Beethoven's 1st and 2nd movements. SC is also significantly higher in Dean's 4th than in Brahms' 1st through 3rd movements. SC was lowest in Brahms's 1st movement overall, but increases over the four movements, with SC in the 3rd and the 4th movements being significantly higher than the 1st and 2nd movements. HR increases over the four movements also in Brahms, with HR significantly higher in the 4th movement than in the 1st and 2nd movements (but still significantly lower than all Beethoven movements). EMG in Dean's 1st movement is significantly higher than Beethoven's 1st, 2nd, and 4th, also Brahms' 1st through 4th movements. EMG in Dean's 4th movement is also significantly higher than Beethoven's 1st and Brahms' 1st and 3rd movements. Dean's 2nd movement presented via loudspeaker did not lead to striking differences in physiology compared to the live played music.

Table 3. Results of the Conditional F-Tests Following the Linear Mixed Effects Models for Each Psychophysiological Measure and the Fixed Effects of Work or Movement

	Work					Movement				
Measure	F	df	*p*	$R^2_{(m)}$	$R^2_{(c)}$	F	df	*p*	$R^2_{(m)}$	$R^2_{(c)}$
EMG	5.54	2, 8.99	.027	.018	0.258	4.559	11, 1,065	<.001	0.032	0.260
RR	1.18	2, 9.01	.35	.002	0.577	2.488	11, 1,066	.004	0.010	0.578
HR	36.88	2, 9.05	<.001	.018	0.946	25.338	11, 700	<.001	0.020	0.947
SC	4.24	2, 8.96	.051	.024	0.633	11.922	11, 871	<.001	0.050	0.632

Relations Between Psychophysiological Measures and Self-Reports

The random forest regressions revealed higher importance values in SC for feeling energetic, joyful, indifferent, powerful, inspired, and agitated, and feelings of tenderness (R^2 = .064; Figure 4), where SC activity increases with feeling energetic, joyful, indifferent, and agitated, and decreases with feeling powerful and inspired and feelings of tenderness. Importance values for HR were high for feeling inspired and forgetting about being in a concert (the situation) and surroundings (R^2 = .138), where HR increases with feeling inspired but decreases when participants report they forgot being in a concert and the surroundings. Higher importance values in RR were seen for feelings of tenderness and feeling indifferent, energetic, and powerful (R^2 = .106), where RR increases with feeling energetic and indifferent but decreases with feelings of tenderness and feeling powerful. In EMG, importance values were high for feelings of tenderness; feeling relaxed, moved, and inspired; and liking (where EMG activity decreases with these feelings; see online supplemental materials), although the explained variance is low (R^2 = .004).

Linear models with the latent variables revealed that SC activity was predicted by mixed and negative emotions (explained variance: R^2(m) = .015, R^2(c) = .61; beta estimates, standard errors, and significance levels are depicted in Figure 5) and dissociation (R^2(m) = .012, R^2(c) = .62), that is, SC activity decreased with higher mixed emotions, and increased with higher negative emotions and dissociation. HR was predicted by positive emotions (R^2(m) = .001, R^2(c) = .94), that is, increased with higher positive emotions. RR was only marginally predicted by mixed emotions (p = .086), that is, decreased with higher mixed emotions. EMG was significantly predicted by negative emotions (R^2(m) = .010, R^2(c) = .28), showing that EMG activity increased with higher negative emotions.

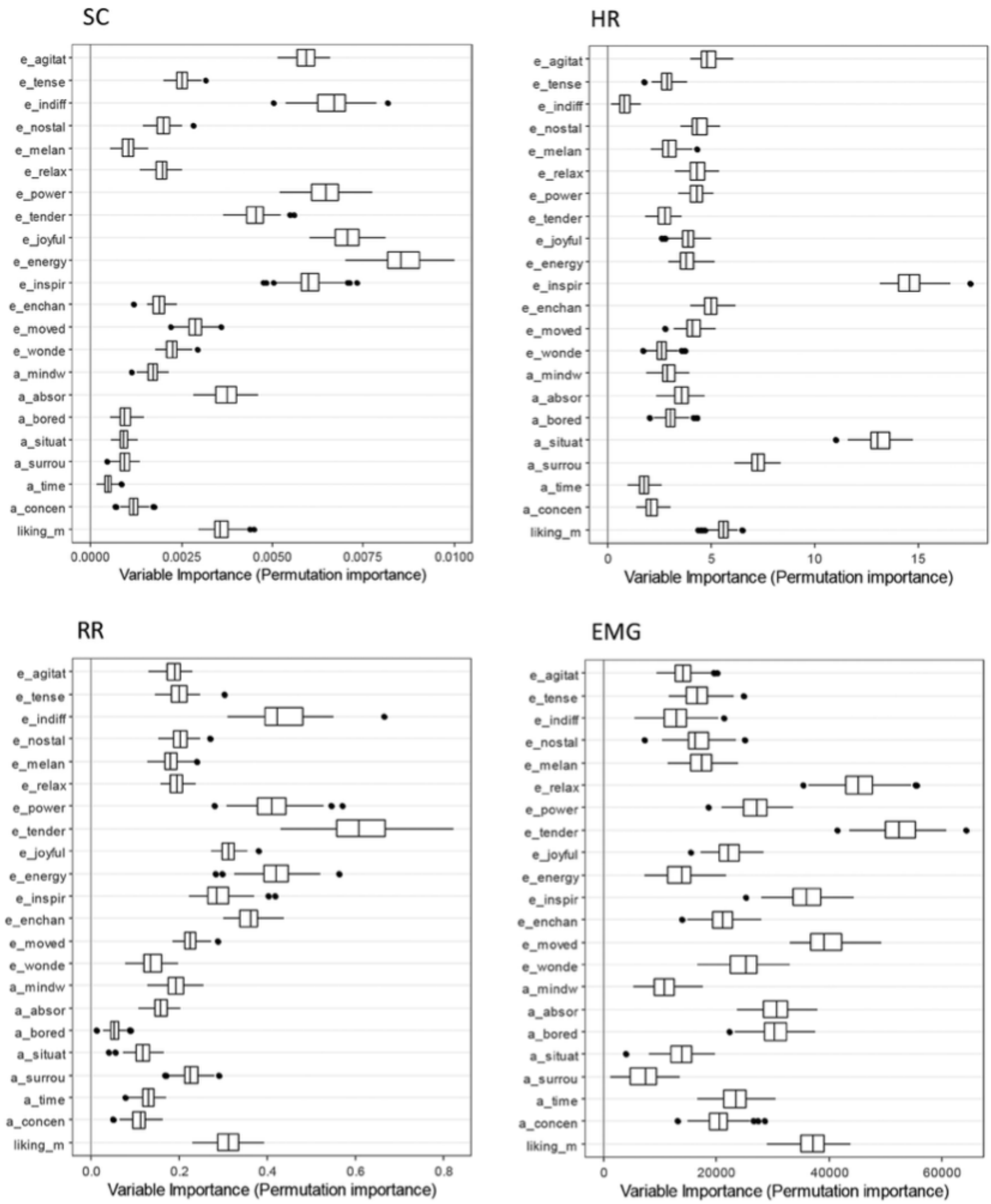


Figure 4. Importance Values for Each Emotion and Absorption Item and Physiological Measure. Note. Variable importance (%IncMSE) = percent increase in mean squared error of predictions (values in their original units); e_ = GEMIAC; a_ = absorption scale; _m = liking rating per movement.

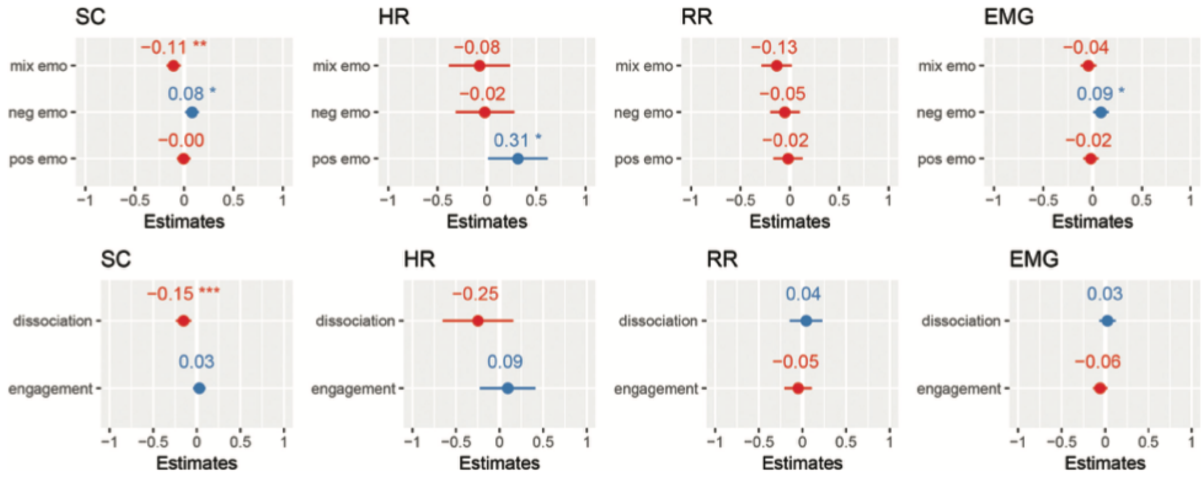


Figure 5. Beta Estimates from the Linear Mixed Effects Models for All Latent Variables. Note. See the online article for the colour version of this figure. * $p < .05$. ** $p < .01$. *** $p < .001$.

Movement	Classical – Beethoven			Section	Romantic – Brahms			Movement
	Physiology	*Emotions Absorption*	*Tonal structure*		*Tonal structure*	*Emotions Absorption*	*Physiology*	
1			tonic	opening	tonic		- HR, - SC	1
2		mixed	nontonic	inner	nontonic	mixed		2
3	(+ EMG)	positive, engagement	tonic		nontonic	mixed		3
4		positive	tonic	closing	tonic	positive, engagement	+ HR, + SC	4

Figure 6. Grouping of Movements in the Classical and the Romantic Works

4 Discussion

This study investigated aesthetic experiences with music in a live concert context, measured by self-reports of music-induced emotions (GEMIAC; Coutinho & Scherer, 2017a) and absorption (Vroegh, 2018), in combination with physiological measures (skin conductance, zygomaticus major muscle activity, heart rate, and respiration rate). We saw how these measures reflect attentive listening in the concert as a dynamic experience (1) on a global level across three works from different musical periods as well as (2) within works on a structural level (between individual movements). We also (3) extend previously found associations of physiological responses with music-evoked emotions as well as positive and negative aesthetic experiences occurring in a real-world concert setting.

The Concert Experience

A concert frame is characterised by presenting a number of works in an intended order, where we observed a rise-and-fall trajectory of emotions (Balteș & Miu, 2014): Positive and mixed emotions started high for the performance of the classical work, dipped during the contemporary work, and rose again for the romantic work. The negative emotions (e.g., tense, agitated) were significantly higher in the contemporary work compared to the classical and romantic works. This shows, on the one hand, that the less familiar work by Dean seemed to be a suitable counterpart to the other two more familiar works, evoking emotions of different valence and arousal (also seen in higher SC and EMG activity) and therefore extending the range of music-evoked emotions in the concert. On the other hand, the work by Dean did not significantly differ in engagement from the other two works, and the concert led to an overall positive experience (high overall liking, averaging at 4.10 out of 5). This finding is relatively important, as it supports the idea of general programming in classical concerts to include new works and placing it between well-known canonic works to increase its chance of being "audience-friendly" (Blackburn, 2016; Gotham, 2014).

In terms of physiology, HR decreased constantly over the course of the concert, from (on average) 72 BPM in Beethoven's 1st movement to 68 BPM in Brahms' 2nd movement. The HR decrease may indicate the general relaxation of the body in the evening (because of a physiological increase in parasympathetic activity) complemented by a calming and relaxing concert experience. Only in the last two movements of the romantic work, HR significantly increased again, but it did not rise to the level of the beginning (i.e., is still significantly lower than Beethoven's 4th movement), potentially reflecting the exciting climactic ending of the Brahms and the awareness of the imminent end of the concert. However, there was the limitation that the order of the works was the same across the concerts. This ecological choice meant that it was not possible to see if this physiological trajectory was evoked by the music (which is relatively unlikely, given that the works exhibited an alternation of movements with different arousal potentials, not a generally decreasing trajectory), the concert frame, or by other unspecific factors.

In a similar way to how the self-report measures were most contrasting in the contemporary work, the physiology in the work by Dean also stood out from the other works: SC in Dean's 1st movement was significantly higher than for the first two movements of both the classical and the romantic works, suggesting that increased skin conductance as evoked by unexpected harmony on a phrase scale (Steinbeis et al., 2006) or single melodic line (Egermann et al., 2013) can be generalised to long-term dissonance across a whole piece. Additionally, the 1st and 4th movements by Dean led to overall significantly higher EMG responses of the zygomaticus muscles (compared to all but Beethoven's 3rd movement). Although inconsistent with positive valence judgments increasing zygomaticus muscle activity (Kirsch et al., 2016), our results are supported by the finding that dissonant (compared to consonant) music elicits significantly more zygomaticus major activity, potentially reflecting "grimacing" or ironic laughter (Dellacherie et al., 2011). The assumption that zygomaticus major activity is not merely involved with smiling (i.e., reflect pleasure, liking, or other positively valenced reactions and evaluations) is further supported by Wingenbach et al. (2020), who recently showed that the zygomaticus major muscle in facial mimicry was also associated with negatively valenced emotions such as fear and disgust, the latter already indicated in a music related setting (Vrana, 1993).

Finally, as it has been theorised that aesthetic experiences in music may be more intensely evoked by a combination of auditory and visual information (Cohen, 2009; Huron, 2012), enhanced in a live performance by adding a sense of creativity and uniqueness to the experience (Auslander, 2002; Brown & Knox, 2017; Holt, 2010; Minor et al., 2004; Packer & Ballantyne, 2011), we expected that the recorded music would evoke less intense positive aesthetic experiences compared to the live music. The current study found that the music presented as a recording – rather than live – was rated the lowest for engagement compared to four other movements. This shows that participants were less concentrated and absorbed and liked it less, probably due to the absence of a main factor of liveness, that is, the possibility to watch the musicians creating the music in the moment (Seibert et al., 2020). Despite being also in line with other studies comparing a live performance with an audio-visual recording (Coutinho & Scherer, 2017b), the comparison of live and recorded music did not reveal striking significant differences in other self-reports or physiological responses. This result gives rise to the assumption that the situation moderated the aesthetic experience, which points to the frame (Goffman, 1974) being a key component in modelling the aesthetic experience (Leder & Nadal, 2014; Locher et al., 2010; Pelowski et al., 2016). This is further supported by the fact that Coutinho and Scherer (2017b) compared responses from a live performance in a typical concert venue (church) and the recorded performance in a university lecture hall, that is, that the difference in emotion may have been due to the lecture hall versus concert frame, rather than necessarily due to the recorded or liveness per se. However, the limitation here is that only one movement was presented as a recording, and only from the contemporary work. This compromise was made to keep a more "naturalistic" approach as only the contemporary style could entail electro-acoustic elements. Nonetheless, this could be manipulated in future studies to confirm the role of engagement in a live concert setting.

Effects of Compositional Trajectory

While the performances of the classical and the romantic works hardly differed as whole works in self-reports and physiological responses, similarities as well as differences between movements were evident in these two genres. This could reflect the individual compositional style and strategy of the composers, and thus the overall organisation of the work.

Based on the self-report data, participants seemed to perceive three sections of these forms of a sonata: an opening section (Beethoven's and Brahms' 1st movements), an inner section (Beethoven's and Brahms' 2nd movements, and Brahms 3rd movements), and a closing section (Beethoven's 3rd movement, Beethoven's and Brahms' 4th movement). The inner sections evoked higher mixed emotions (lower arousal emotions such as being moved, nostalgic, melancholic) and lower positive emotions, while the closing sections evoked lower mixed and higher positive emotions (such as energetic, joyful, inspired) as well as higher engagement. This reflection of the formal structure of the works is depicted in Figure 6.

In a similar way that storytelling is perceived in a linear trajectory (Simony et al., 2016), Webster describes a musical "plot" (Webster, 1991, p. 182) and analytically differentiates between the complementary perspectives of cyclic integration (thematicism, tonal organisation, p. 194) and the so-called through-composition. The idea of the through-composition focuses on aspects of gestural phenomena and rhetoric (p. 123), which imply an audience that is capable through anticipation, expectation, and involvement to conceive the work as a whole. (Czerny, 1848/1930, pp. 33–81) has already suggested that the composition of a sonata cycle incorporates the concept of unity within and throughout its movements, by means of compatibility of affect (tempo, character) and unity of keys (see also Ratner, 1980, p. 322). Within a sonata form, the question arises of how listeners perceive the overall trajectory of the four movements. Hepokoski and Darcy (2006), for example, proposed a binary-based (2 + 2) structure, which is mainly based on the symmetrical tonal development from the tonic to a nontonic key (movements one and two) and its resolving return to the tonic (movements three and four), indicating a higher contrast between the 2nd and the 3rd movement.

The current results can relate to this musical theoretical discourse, albeit more related to the emotional content of the movements (Balteș & Miu, 2014): in the classical and romantic works used in this study, there are sections of contrasting emotion/arousal, forming a third section emphasising the nontonic space (associated with more tension) of the inner movements between the outer movements in the tonic key. This is in line with Tovey (1956), who has described the “emotional content” of a sonata cycle and observed that the individual movements raised emotional issues that cannot be resolved without the other movements (p. 230). These results show that the reported felt emotions reflect the character/function of the movements, related to previous work showing matched relationships between felt and perceived emotions (Evans & Schubert, 2008).

Evidence for the different groupings between the classical and romantic works can also be seen in physiology, specifically in the 3rd movements of both works. While the four movements of the classical work do not show much physiological variation, the last two movements from the romantic work show a significant increase in HR and SC, probably reflecting the structure of the romantic work, in which the normative key-order scheme is changed: The third movement is not yet in the tonic key, thus serving as a preparation of the climax and helping to establish a clear area of return and relaxation (tonic key) only with the very last movement (Hepokoski & Darcy, 2006, p. 340). This interplay of tension (nontonic) and relaxation (tonic) is a typical stylistic feature of tonal music of the common practice period and affects general expectation and prediction in listeners on all levels (Huron, 2006; Sears et al., 2019).

These responses within the sonata form specifically occur in the exemplars of the common practice period, but were not present in the contemporary work by Dean. As expected, the linear trajectory of the “suite” led to a very different sequence of emotion and tension ratings. Here, an increase in positive and a decrease in negative emotions from the 1st to the 4th movements of the contemporary work was observed, suggesting that liking may increase with familiarity in musical style (Szpunar et al., 2004) even in complex music (Madison & Schiölde, 2017). Again, the ratings fit to the overall trajectory of the composition, which cannot be elaborated further in this context. Nonetheless, it has to be noted that contemporary music also plays with expectation, tension, and relaxation, but in a more diverse way, and it is very well possible for future studies to draw on schemata that are less common compared to the sonata form, and also in the contemporary style. Taken together, the (music-)theoretical approaches fit with the findings of this study and suggest that there might be at least an implicit perception of inner coherence between movements within the sonata form.

Relations Between Self-Reports and Physiology in an Aesthetic Context

By operationalizing a musical aesthetic experience with emotions and absorption (Brattico et al., 2013; Brattico & Pearce, 2013), this study investigated the aesthetic experience of not only positive (e.g., Salimpoor et al., 2009), but negative and mixed-valence emotions in response to contemporary, classical, and romantic music performances in combination with continuously acquired psychophysiological measures. Therefore, our analysis was somewhat exploratory, and we conducted two separate analyses with two kinds of predictors for the physiological responses: the single items from each scale and the latent variables obtained from the factor analysis (independent of a specific musical work).

The models from the random forest regressions with single self- report items from the GEMIAC and absorption scales show small to medium effects, while linear mixed effects models with the estimated five latent variables predicting physiology did not reveal noteworthy effects (all $R^2 < .02$). This means that the results from the random forest regression (i.e., single items) were more successful in explaining the variance in the physiology measures. These results could be due to the fact that latent factors were grouped by valence (where positive and negative emotion factors were both high arousal), whereas physiology tends to also respond to differences in arousal (Baumgartner et al., 2006; Bradley & Lang, 2000; Egermann et al., 2013; Gomez & Danuser, 2004). Indeed, our current results support these studies, by showing the strongest relations with typical arousal reactions, for example, SC and RR increased with high arousal emotions of feeling energetic and agitated, while SC also increased with joyfulness. SC, RR, and EMG responses decreased with lower arousal emotions such as feelings of tenderness.

Another reason why the single items predicted the physiology better than the latent variables could be because the music in the current study was not chosen to represent one "basic" emotion (Baumgartner et al., 2006; Etzel et al., 2006; Khalfa et al., 2002; Krumhansl, 1997; Lundqvist et al., 2008) or separate locations on the arousal and valence dimensions (Dillman Carpentier & Potter, 2007; Egermann et al., 2015; Gomez & Danuser, 2004). Rather, they were chosen for their potential to evoke diverse and dynamic aesthetic experiences due to their quality as being widely appreciated compositions (for at least two of the works). Therefore, music-induced emotions that cannot be expressed by the music – such as feeling inspired, moved, or even bored – were investigated.

The results reveal that feeling inspired – considered a positive emotion – was associated with a decrease in EMG and SC, but an increase in HR, supporting previous research that heart rate is positively related to the experienced valence of the music (Sammler et al., 2007). Feeling powerful was associated with decreased SC and RR, suggesting that it may be related to a calm state of mind. Interestingly, we found that feeling indifferent/bored was associated with increased SC and RR. We also found that being moved and liking ratings were associated with a decrease of smiling muscle activity, which somewhat contradicts what we were expecting (Bradley & Lang, 2000; Kirsch et al., 2016). An explanation for the current findings might be that idea that more positive musical judgments are associated with motor inhibition indices, such as smaller eyeblink amplitudes (Roy et al., 2009), slower response times, and a higher amplitude of a motor inhibition N2/P3 component in an electroencephalography (EEG) event-related potential brain response (Sarasso et al., 2019). According to this perspective, positive aesthetic emotions and judgements inhibit muscle activity, maximising attentional resources on the perception of the aesthetic object. However, as mentioned above, the negative relationship with EMG activity of the smiling muscle and positive feelings/liking also could be due to the activation of EMG muscles reflecting a negative experience, such as a grimace (e.g., Dellacherie et al., 2011; Wingenbach et al., 2020). As we only took measurements from one facial EMG muscle, our current findings are not conclusive as to whether a negative experience is the result of an inhibition of this muscle during more positive emotions or an activation in response to negative felt emotions, but it could be possible that both mechanisms were at play here.

The physiology seemed to reflect two absorption items as well: HR decreased with the items forgetting about being in a concert and the surroundings, which is in line with the conclusion (made above) that the overall concert experience is reflected in HR – mirrored in items of dissociation, that is, the exclusion of other content from the phenomenal field (Butler, 2004; Vroegh, 2018).

Although we show how physiology may be linked to emotions (on general arousal-valence dimensions) and absorption, a limitation of these measures is that they do not directly index emotions without being associated with self-reports of emotions and feelings. Hence, this study is an attempt at formulating more general conclusions on which emotions and physiological responses occur during a classical concert in response to different kinds of music. Further research may need to reduce this set of emotions and include continuous ratings to track the relations between continuous changes in rating behaviour and physiology.

5 Conclusion

This exploratory study considered the aesthetic experience – as operationalized by music-induced emotions and absorption – in combination with physiology across whole multi-movement works of music of varying styles framed in a typical classical concert setting. Conducting research in a frame where the music is placed for optimum attention (e.g., dimmed lights, expected total silence), we observed, first, (a) a rise-and-fall trajectory of emotions – with the contemporary work leading to higher negatively valenced emotions – embedded in an overall highly liked and appreciated concert and (b) that the movement presented as a recording had lower engagement than movements performed live. These results tentatively highlight the programming and "liveness" as components of the concert frame which may influence the aesthetic experience – though it remains to be more systematically manipulated in future studies to confirm this effect. Second, investigating the experience of whole works yielded novel insights into the perception of a (cyclic) multi-movement work. Emotional and – in part – physiological responses mirrored the functions across movements in familiar classical and romantic works representing variations of the sonata form, contrasting with the unfamiliar contemporary work (in the form of a suite), which did not follow such a trajectory. Third, by selecting naturalistic music for its potential to evoke diverse aesthetic experiences, we extended previous results by showing that physiological responses reflect a range of music- evoked emotions relating to both positive and negative aesthetic experiences. To elaborate on these findings, it seems exploratory research is required in combination with more controlled studies to further understand the interplay between compositional conventions and musical expectations, between the score and the act of performance in a specific frame, and the resulting live concert experience (Toelle & Sloboda, 2021; Wald-Fuhrmann et al., 2021).

6 Supplementary Materials

Methods

Participants

Concert visits in general. Participants were asked which kinds of concerts they mostly visit, multiple answers were possible: Most visited concerts comprise classical concerts (N=19, 25, 12) and Opera (N=6, 21, 8), in C1, also Rock and Pop concerts are named, < 10 are Jazz and contemporary music, church music and club/disco, musical.

Company. Half of the participants came to the concerts in company: C1 (N=18), C2 (N=17), C3 (N=11).

After each piece (i.e., Beethoven, Dean and Brahms), participants were additionally asked, if they liked the piece as a whole, if they liked the performance, if they were familiar with this kind of music, if they wanted to move to the music, if they wanted to 'understand' the music, if they felt a connection with the musicians and with the other listeners (all on a 5-point Likert scale from not at all (1) to very much (5)). Finally, if they knew the piece (yes, no, I am not sure).

After the concert (i.e., the last piece), they were additionally asked, if they liked the concert as a whole, if they had the feeling to actually visit a concert (rather than taking part in a study), if they liked the atmosphere of the concert, if they were bothered by the electrodes, if they thought the preparation time was adequate, if the duration of the concert was adequate and finally, if they could imagine, visiting such a research concert again.

Distribution of ratings

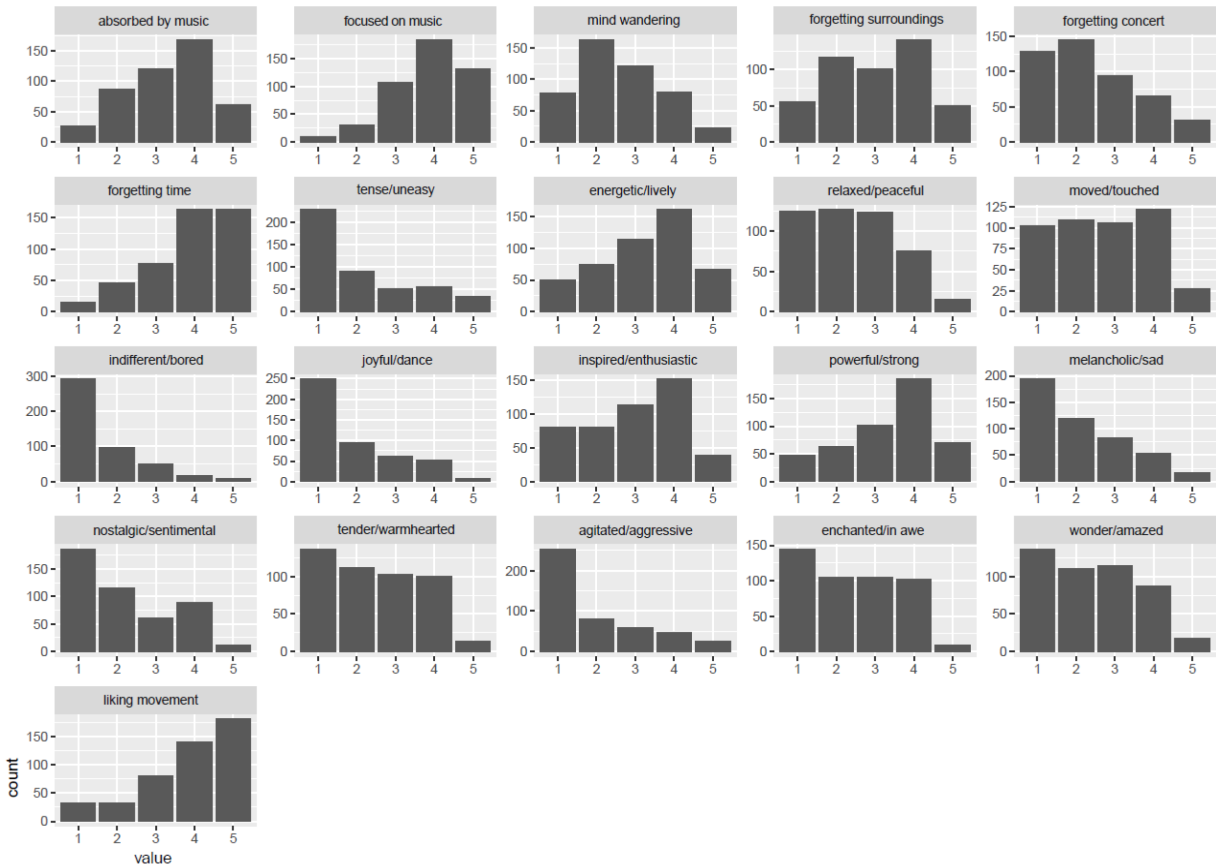


Figure S1. Histograms of all items. Ratings from not at all (1) to very much (5).

Concert comparison of psychophysiology

Results of the conditional *F*-tests for Concert following the linear mixed models:

- SC: $F(2, 84.7) = 0.76$, $p = .47$
- HR: $F(2, 79) = 0.11$, $p = .896$
- RR: $F(2, 95) = 0.091$, $p = .914$
- EMG: $F(2, 95) = 1.065$, $p = .349$

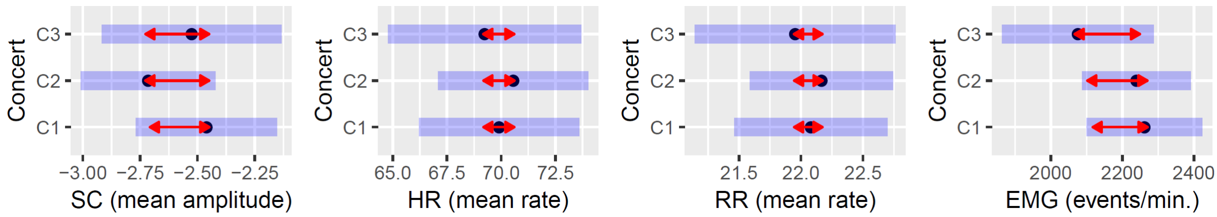


Figure S2. Comparisons of psychophysiological measures between concerts. Mean of measure with confidence intervals (blue bars) and results of the pairwise comparisons with emmeans (red arrows). If red arrows do not overlap, there is a significant difference between conditions.

Effects of piece and movement on self-reports (latent variables)

	pos emo			mix emo			neg emo			engagement			dissociation		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
(Intercept)	0.19	-0.03 – 0.40	0.085	0.13	-0.05 – 0.31	0.167	-0.23	-0.41 – -0.05	**0.013**	0.04	-0.18 – 0.25	0.722	-0.06	-0.26 – 0.14	0.548
1.2 Beethoven	-0.31	-0.55 – -0.07	**0.013**	0.40	0.19 – 0.60	**<0.001**	-0.35	-0.56 – -0.14	**0.001**	-0.18	-0.44 – 0.07	0.160	0.03	-0.18 – 0.25	0.749
1.3 Beethoven	0.19	-0.05 – 0.43	0.117	-0.22	-0.42 – -0.03	**0.028**	-0.19	-0.40 – 0.02	0.077	0.23	-0.02 – 0.48	0.071	0.05	-0.16 – 0.26	0.629
1.4 Beethoven	0.31	0.07 – 0.54	**0.012**	-0.26	-0.46 – -0.06	**0.011**	-0.02	-0.23 – 0.19	0.876	0.16	-0.09 – 0.41	0.202	0.05	-0.15 – 0.26	0.609
2.1 Dean	-0.79	-1.03 – -0.55	**<0.001**	-0.43	-0.63 – -0.23	**<0.001**	1.06	0.85 – 1.27	**<0.001**	-0.16	-0.41 – 0.10	0.226	-0.05	-0.26 – 0.16	0.616
2.2 Dean	-0.62	-0.86 – -0.38	**<0.001**	-0.63	-0.83 – -0.43	**<0.001**	1.12	0.91 – 1.33	**<0.001**	-0.38	-0.63 – -0.13	**0.003**	-0.09	-0.30 – 0.11	0.376
2.3 Dean	-0.47	-0.71 – -0.23	**<0.001**	-0.56	-0.76 – -0.35	**<0.001**	1.00	0.79 – 1.22	**<0.001**	-0.02	-0.27 – 0.23	0.861	0.09	-0.12 – 0.30	0.384
2.4 Dean	-0.15	-0.38 – 0.09	0.218	-0.77	-0.96 – -0.57	**<0.001**	0.70	0.49 – 0.91	**<0.001**	0.04	-0.21 – 0.28	0.776	-0.12	-0.33 – 0.08	0.238
3.1 Brahms	-0.11	-0.35 – 0.14	0.390	-0.03	-0.24 – 0.17	0.748	-0.09	-0.31 – 0.12	0.393	-0.17	-0.42 – 0.09	0.195	0.11	-0.10 – 0.32	0.299
3.2 Brahms	-0.35	-0.58 – -0.11	**0.005**	0.89	0.69 – 1.09	**<0.001**	-0.26	-0.47 – -0.05	**0.015**	-0.08	-0.33 – 0.17	0.522	0.35	0.14 – 0.56	**0.001**
3.3 Brahms	-0.41	-0.65 – -0.17	**0.001**	0.47	0.27 – 0.67	**<0.001**	-0.34	-0.56 – -0.13	**0.002**	-0.14	-0.40 – 0.11	0.279	0.20	-0.02 – 0.41	0.070
3.4 Brahms	0.31	0.07 – 0.55	**0.011**	-0.27	-0.47 – -0.07	**0.008**	-0.15	-0.36 – 0.06	0.153	0.31	0.06 – 0.56	**0.015**	0.08	-0.12 – 0.29	0.424
Random Effects															
σ^2	0.50			0.35			0.39			0.56			0.38		
τ_{00}	0.29 Participant			0.25 Participant			0.17 Participant			0.25 Participant			0.34 Participant		
ICC	0.36			0.42			0.30			0.31			0.47		
N	88 Participant			88 Participant			88 Participant			88 Participant			88 Participant		
Observations	909			909			909			909			909		
Marginal R^2 / Conditional R^2	0.131 / 0.447			0.275 / 0.576			0.360 / 0.551			0.043 / 0.338			0.021 / 0.479		

	pos emo			mix emo			neg emo			engagement			dissociation		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
(Intercept)	0.23	-0.07 – 0.54	0.161	0.11	-0.26 – 0.47	0.578	-0.37	-0.55 – -0.19	**0.001**	0.09	-0.13 – 0.32	0.422	-0.02	-0.18 – 0.13	0.754
2 Dean	-0.55	-0.96 – -0.15	**0.024**	-0.57	-1.07 – -0.08	**0.048**	1.11	0.89 – 1.33	**<0.001**	-0.18	-0.46 – 0.09	0.223	-0.08	-0.21 – 0.04	0.234
3 Brahms	-0.19	-0.59 – 0.22	0.390	0.29	-0.21 – 0.78	0.285	-0.07	-0.29 – 0.15	0.531	-0.07	-0.35 – 0.20	0.617	0.15	0.02 – 0.28	**0.045**
Random Effects															
σ^2	0.50			0.35			0.39			0.56			0.38		
τ_{00}	0.29 Participant			0.25 Participant			0.17 Participant			0.25 Participant			0.34 Participant		
	0.08 Movement			0.12 Movement			0.02 Movement			0.03 Movement			0.00 Movement		
ICC	0.42			0.51			0.32			0.34			0.47		
N	12 Movement			12 Movement			12 Movement			12 Movement			12 Movement		
	88 Participant			88 Participant			88 Participant			88 Participant			88 Participant		
Observations	909			909			909			909			909		
Marginal R^2 / Conditional R^2	0.058 / 0.454			0.153 / 0.589			0.339 / 0.552			0.007 / 0.340			0.013 / 0.477		

Effects of concert, piece, and movement on psychophysiology

	EMG			RR			HR			Phasic		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
(Intercept)	2261.56	2101.70 – 2421.41	**<0.001**	22.08	21.47 – 22.69	**<0.001**	69.90	66.27 – 73.54	**<0.001**	-2.46	-2.76 – -2.16	**<0.001**
C 2	-22.26	-241.38 – 196.86	0.843	0.09	-0.75 – 0.92	0.841	0.65	-4.34 – 5.64	0.800	-0.25	-0.67 – 0.16	0.236
C 3	-186.06	-449.41 – 77.30	0.169	-0.13	-1.13 – 0.88	0.806	-0.67	-6.37 – 5.03	0.818	-0.07	-0.56 – 0.43	0.795
Random Effects												
σ^2	633800.01			2.42			7.81			0.53		
τ_{00}	186648.09 Participant			3.28 Participant			99.03 Participant			0.72 Participant		
ICC	0.23			0.57			0.93			0.58		
N	98 Participant			98 Participant			82 Participant			88 Participant		
Observations	1174			1175			793			969		
Marginal R^2 / Conditional R^2	0.006 / 0.232			0.001 / 0.575			0.002 / 0.927			0.011 / 0.581		

	EMG			RR			HR			Phasic		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
(Intercept)	2164.85	2015.14 – 2314.56	**<0.001**	21.96	21.53 – 22.39	**<0.001**	71.73	69.52 – 73.94	**<0.001**	-2.52	-2.79 – -2.25	**<0.001**
2 Dean	212.81	39.50 – 386.13	**0.039**	0.26	-0.07 – 0.60	0.159	-2.17	-2.93 – -1.41	**<0.001**	0.11	-0.18 – 0.40	0.485
3 Brahms	-69.63	-242.90 – 103.64	0.451	0.12	-0.21 – 0.46	0.491	-3.30	-4.06 – -2.53	**<0.001**	-0.31	-0.60 – -0.02	0.067
Random Effects												
σ^2	611599.19			2.39			5.67			0.46		
τ_{00}	188614.92 Participant			3.21 Participant			98.14 Participant			0.73 Participant		
	9381.71 Movement			0.03 Movement			0.21 Movement			0.04 Movement		
ICC	0.24			0.58			0.95			0.62		
N	12 Movement			12 Movement			12 Movement			12 Movement		
	98 Participant			98 Participant			82 Participant			88 Participant		
Observations	1174			1175			793			969		
Marginal R^2 / Conditional R^2	0.018 / 0.258			0.002 / 0.577			0.018 / 0.946			0.024 / 0.633		

	EMG			RR			HR			Phasic		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
(Intercept)	1993.12	1816.01 – 2170.22	**<0.001**	21.82	21.35 – 22.29	**<0.001**	72.42	70.20 – 74.64	**<0.001**	-2.59	-2.82 – -2.36	**<0.001**
1.2 Beethoven	184.26	-34.71 – 403.24	0.099	0.28	-0.16 – 0.71	0.212	-0.58	-1.38 – 0.22	0.153	-0.03	-0.23 – 0.18	0.801
1.3 Beethoven	323.50	104.53 – 542.47	**0.004**	0.14	-0.30 – 0.57	0.532	-1.10	-1.90 – -0.31	**0.007**	0.19	-0.02 – 0.40	0.071
1.4 Beethoven	179.27	-40.31 – 398.85	0.110	0.15	-0.28 – 0.58	0.501	-1.09	-1.91 – -0.28	**0.008**	0.11	-0.10 – 0.31	0.308
2.1 Dean	565.01	345.44 – 784.59	**<0.001**	0.68	0.25 – 1.11	**0.002**	-2.58	-3.39 – -1.77	**<0.001**	0.44	0.24 – 0.65	**<0.001**
2.2 Dean	269.83	50.86 – 488.81	**0.016**	0.26	-0.17 – 0.70	0.233	-2.81	-3.61 – -2.01	**<0.001**	0.07	-0.14 – 0.27	0.533
2.3 Dean	235.60	16.63 – 454.57	**0.035**	0.09	-0.34 – 0.53	0.669	-3.21	-4.01 – -2.41	**<0.001**	-0.04	-0.25 – 0.17	0.710
2.4 Dean	468.55	249.58 – 687.52	**<0.001**	0.58	0.15 – 1.01	**0.009**	-2.86	-3.66 – -2.05	**<0.001**	0.23	0.02 – 0.44	**0.029**
3.1 Brahms	47.11	-171.86 – 266.08	0.673	0.10	-0.33 – 0.53	0.648	-4.58	-5.42 – -3.74	**<0.001**	-0.53	-0.73 – -0.32	**<0.001**
3.2 Brahms	117.29	-101.68 – 336.27	0.294	0.35	-0.09 – 0.78	0.117	-4.67	-5.48 – -3.85	**<0.001**	-0.40	-0.61 – -0.19	**<0.001**
3.3 Brahms	54.82	-164.15 – 273.79	0.624	-0.04	-0.47 – 0.40	0.869	-3.60	-4.43 – -2.78	**<0.001**	-0.13	-0.35 – 0.08	0.218
3.4 Brahms	189.21	-29.76 – 408.18	0.091	0.64	0.21 – 1.08	**0.004**	-3.10	-3.94 – -2.27	**<0.001**	0.10	-0.11 – 0.31	0.358
Random Effects												
σ^2	611614.22			2.39			5.67			0.46		
τ_{00}	188542.58 Participant			3.21 Participant			98.26 Participant			0.73 Participant		
ICC	0.24			0.57			0.95			0.61		
N	98 Participant			98 Participant			82 Participant			88 Participant		
Observations	1174			1175			793			969		
Marginal R^2 / Conditional R^2	0.032 / 0.260			0.010 / 0.578			0.020 / 0.947			0.050 / 0.632		

Physiology and self-report data (latent variables)

	EMG			Phasic			HR			RR		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
(Intercept)	0.02	-0.11 – 0.14	0.757	-2.59	-2.78 – -2.40	**<0.001**	69.92	67.64 – 72.20	**<0.001**	22.07	21.67 – 22.48	**<0.001**
pos emo	-0.02	-0.10 – 0.06	0.665	-0.00	-0.07 – 0.06	0.891	0.31	0.01 – 0.62	**0.043**	-0.02	-0.17 – 0.13	0.805
neg emo	0.09	0.01 – 0.16	**0.031**	0.08	0.01 – 0.15	**0.023**	-0.02	-0.32 – 0.28	0.889	-0.05	-0.20 – 0.10	0.525
mix emo	-0.04	-0.12 – 0.04	0.299	-0.11	-0.18 – -0.04	**0.004**	-0.08	-0.39 – 0.23	0.629	-0.13	-0.29 – 0.02	0.086
Random Effects												
σ^2	0.76			0.48			6.85			2.54		
τ_{00}	0.28 Participant			0.73 Participant			101.36 Participant			3.49 Participant		
ICC	0.27			0.60			0.94			0.58		
N	88 Participant			85 Participant			76 Participant			88 Participant		
Observations	907			808			624			908		
Marginal R^2 / Conditional R^2	0.010 / 0.278			0.015 / 0.609			0.001 / 0.937			0.003 / 0.580		

	EMG			Phasic			HR			RR		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
(Intercept)	0.02	-0.10 – 0.14	0.752	-2.59	-2.78 – -2.39	**<0.001**	69.90	67.63 – 72.18	**<0.001**	22.07	21.67 – 22.48	**<0.001**
engagement	-0.06	-0.14 – 0.02	0.169	0.03	-0.04 – 0.10	0.433	0.09	-0.23 – 0.41	0.570	-0.05	-0.21 – 0.11	0.524
dissociation	0.03	-0.07 – 0.12	0.596	-0.15	-0.24 – -0.06	**0.001**	-0.25	-0.65 – 0.15	0.226	0.04	-0.15 – 0.23	0.664
Random Effects												
σ^2	0.77			0.49			6.88			2.55		
τ_{00}	0.26 Participant			0.77 Participant			101.04 Participant			3.46 Participant		
ICC	0.25			0.61			0.94			0.58		
N	88 Participant			85 Participant			76 Participant			88 Participant		
Observations	907			808			624			908		
Marginal R^2 / Conditional R^2	0.002 / 0.255			0.012 / 0.617			0.000 / 0.936			0.000 / 0.575		

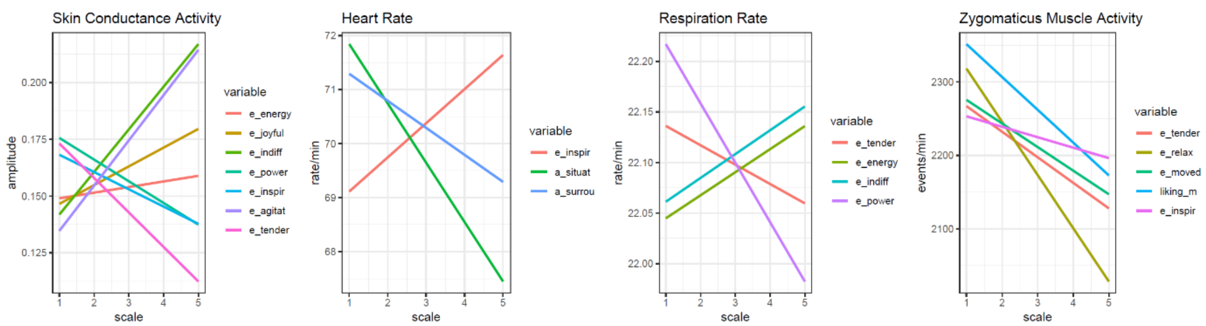


Figure S3. Psychophysiological measures in relation to emotion and absorption items that show higher importance values based on random forest regressions.

Chapter 3

Inter-subject correlation of physiological responses during live music concerts

Based on:

Czepiel, A., Fink, L. K., Fink, L. T., Wald-Fuhrmann, M., Tröndle, M., & Merrill, J. (2021). Synchrony in the periphery: inter-subject correlation of physiological responses during live music concerts. *Scientific reports*, *11*(1), 1-16. https://doi.org/10.1038/s41598-021-00492-3

Abstract

While there is an increasing shift in cognitive science to study perception of naturalistic stimuli, this study extends this goal to naturalistic contexts by assessing physiological synchrony across audience members in a concert setting. Cardiorespiratory, skin conductance, and facial muscle responses were measured from participants attending live string quintet performances of full-length works from Viennese Classical, Contemporary, and Romantic styles. The concert was repeated on three consecutive days with different audiences. Using inter-subject correlation (ISC) to identify reliable responses to music, we found that highly correlated responses depicted typical signatures of physiological arousal. By relating physiological ISC to quantitative values of music features, logistic regressions revealed that high physiological synchrony was consistently predicted by faster tempi (which had higher ratings of arousing emotions and engagement), but only in Classical and Romantic styles (rated as familiar) and not the Contemporary style (rated as unfamiliar). Additionally, highly synchronised responses across all three concert audiences occurred during important structural moments in the music – identified using music theoretical analysis – namely at transitional passages, boundaries, and phrase repetitions. Overall, our results show that specific music features induce similar physiological responses across audience members in a concert context, which are linked to arousal, engagement, and familiarity.

1 Introduction

While there is an increasing shift in cognitive science to study human perception of naturalistic stimuli (e.g., real-world movies or music, Nastase et al., 2020), such research is still required in more naturalistic contexts. A concert setting provides one promising context to which research focusing on perception and experience of music can be extended; not only does it afford one possible naturalistic setting for music listening, but also live performances can evoke stronger emotional responses (Coutinho & Scherer, 2017b; Gabrielsson & Wik, 2003; Lamont, 2011) and offer a more immersive experience (Phillips et al., 2020; Wald-Fuhrmann et al., 2021). Although brain imaging techniques can implicitly measure undisturbed (i.e., without behavioural ratings) naturalistic musical perception as it evolves (Abrams et al., 2013; Alluri et al., 2012; Madsen et al., 2019), these methods lack applicability in a wider range of typical listening situations; therefore, more portable methods for measuring continuous responses such as motion capture (Swarbrick et al., 2019) or mobile measurement of peripheral physiology (Ardizzi et al., 2020; Egermann et al., 2013) are required. As our interest lay in the musical experience within the context of a Western art music concert – in which listeners are typically still (Thorau & Ziemer, 2019; Wald-Fuhrmann et al., 2021) – we focused on physiological responses of the autonomic system (ANS).

Previous research shows that certain physiological signatures indicative of a momentary ANS activations, regarding (phasic) skin conductance response (SCR, i.e., sweat secretion), heart rate (HR), and respiration rate (RR) as well as responses of facial muscles (electromyography [EMG] measurement), are related to orientation responses (Benedek & Kaernbach, 2010; Frith & Allen, 1983) and affective processing (Bradley, 2009; Bradley & Lang, 2000). Event-related changes in SCR, HR, RR, and EMG – reflecting a startle (Brown et al., 1991; Dimberg, 1990; Graham & Clifton, 1966) or orienting response (Barry & Sokolov, 1993) – have been associated with pitch changes (Lyytinen et al., 1992; Sidle & Heron, 1976) and tone loudness (the louder the sound, the greater the SCR amplitude (Barry, 1975; Chuen et al., 2016), as well as deviations in timbre, rhythm, and tempo (Chuen et al., 2016). In other words, such signatures may be a response to novelty in stimuli.

Additionally, physiological responses have been shown to occur in response to arousing acoustic features (Bradley, 2009; Bradley & Lang, 2000; Cacioppo et al., 2000; Hodges, 2010). For example, faster and increasing tempi are associated with greater arousal (reflecting emotions such as happiness) in the music (Coutinho & Cangelosi, 2011; Gabrielsson, 2012; Gabrielsson & Juslin, 1996; Gomez & Danuser, 2007), which correspond to increased HR (Bernardi et al., 2006; Dillman Carpentier & Potter, 2007; Egermann et al., 2015; Gomez & Danuser, 2007), RR (Krumhansl, 1997), and SCR (Chuen et al., 2016;

Egermann et al., 2015; Gomez & Danuser, 2007; Khalfa et al., 2002; van der Zwaag et al., 2011) in a listener. Slower-paced music (reflecting low arousal emotions such as sadness) reduces HR, RR, and SCR (Etzel et al., 2006; Gupta & Gupta, 2015; Krumhansl, 1997). Timbral features, such as brighter tones and higher spectral centroid, are associated with higher arousal (Brattico et al., 2011; Coutinho & Cangelosi, 2011; Gingras et al., 2014), which correlates somewhat to RR and SCR (Bannister & Eerola, 2018; Egermann et al., 2015; Gorzelańczyk et al., 2017). Loudness is positively correlated with arousal (Coutinho & Cangelosi, 2011; Juslin & Laukka, 2003; Laurier et al., 2009), and, correspondingly, changes in SCR (Bach et al., 2009; Olsen & Stevens, 2013) and HR (Coutinho & Cangelosi, 2011; Wilson & Aiken, 1977). Ambiguous harmony – operationalised as unexpected harmonic chords (i.e., out-of-key chords in place of tonic chords) and unpredictable notes (i.e., surprising notes within a predictable melodic sequence) – as well as dissonance may also be perceived as tense/arousing (Egermann et al., 2013; Koelsch et al., 2008) and lead to event-related increases in SCR (Dellacherie et al., 2011; Egermann et al., 2013; Steinbeis et al., 2006). Similarly, new or unprepared harmony, enharmonic changes, and harmonic acceleration to cadences have been found to evoke chills (Sloboda, 1991), which are also related to increased SCR, HR, and RR when listening to music (Blood & Zatorre, 2001; Craig, 2005; Grewe et al., 2009; Guhn et al., 2007; Rickard, 2004; Salimpoor et al., 2009). Overall, this shows that peripheral physiology is related to arousing acoustic features, namely faster tempos, harmonic ambiguity, loudness, and (to some extent) timbral brightness. Importantly, it seems that increases in self-reported arousal and physiological measures are time-locked together in an event-related fashion (Egermann et al., 2013; Grewe et al., 2007), suggesting that an increase in reported arousal is simultaneously reflected by increased ANS responses.

It is also worth noting that some ANS responses may be modulated by musical style. Previous studies found that HR increases with faster tempo in Classical music but decreases with faster tempo in rock music (Dillman Carpentier & Potter, 2007). HR is overall lower in atonal, compared to tonal music, independent of the emotional characteristics of the music (Proverbio et al., 2015).

Although previous research generally supports the idea that specific physiological features are associated with specific musical features and styles, some of these studies (for reasons of experimental control) have carefully chosen and cut or constructed stimuli to have little variability in acoustic features (e.g., they use a constant tempo and normalise loudness). However, more research into full-length naturalistic stimuli – which typically have a rich dynamic variation of interdependent features – is required (Nastase et al., 2020).

While previous work using naturalistic music has correlated neural and physiological responses to dynamically changing acoustic features (Alluri et al., 2012, 2013; Bannister & Eerola, 2018; Burunat et al., 2016), or extracted epochs based on information content in the music (Egermann et al., 2013), perhaps a more robust way to identify systematic responses to naturalistic stimuli is via synchrony of responses (Ardizzi et al., 2020; Bernardi et al., 2017; Bonetti et al., 2020; Muszynski et al., 2018; Tschacher et al., 2021), in particular via inter-subject correlation (ISC, see review Nastase et al., 2019). This method – in which (neural) responses are correlated across participants exposed to naturalistic stimuli (Hasson et al., 2004) – assumes that signals not related to processing stimuli would not be correlated. Such synchrony research has demonstrated that highly similar responses occur across subjects when exposed to naturalistic films (Bacha-Trams et al., 2020; Dmochowski et al., 2012; Golland et al., 2015; Hasson et al., 2004; Kauppi, 2010), spoken dialogue (Kang & Wheatley, 2017; Schmälzle et al., 2015; Wilson et al., 2008) and text (Ames et al., 2015; Simony et al., 2016), dance (Herbec et al., 2015), and music (Abrams et al., 2013; Alluri et al., 2012; Kaneshiro et al., 2020; Madsen et al., 2019), strongly suggesting that highly reliable and time-locked responses can be evoked by (seemingly uncontrolled) complex stimuli (for a review see Hasson et al., 2010). Although ISC in functional magnetic resonance imaging (fMRI) studies can identify regions of interest (ROIs) for further analysis (Hasson et al., 2004), ISC can also assess which kind of feature(s) within dynamically evolving stimuli evoke highly correlated responses.

In response to auditory stimuli, higher synchronisation (operationalized via ISC) of participants' responses has been associated with emotional arousal, structural coherence, and familiarity. In terms of emotional arousal, higher correlation coefficients of electroencephalography (EEG) (Dmochowski et al., 2012), SCR, and respiration (Bracken et al., 2014) (as well as higher hemodynamic activity in ROIs that had highly correlated activity between participants' fMRI (Hasson et al., 2010) coincided with moments of high arousal in films, such as a close-up of a revolver (Dmochowski et al., 2012), gun-shots or explosions (Hasson et al., 2004), as well as close-ups of faces and emotional shakiness in voice (Bracken et al., 2014). Additionally, physiological synchrony between co-present audience members of movie viewings and theatre performance correlate with convergence of emotional responses and evaluation (Ardizzi et al., 2020; Golland et al., 2015). Regarding structural coherence, ISC is higher when listening to original, compared to phase- scrambled, versions of music (Abrams et al., 2013; Kaneshiro et al., 2020) and spoken text (Simony et al., 2016). Cardiovascular and respiratory synchrony (calculated here by Generalised Partial Directed Coherence) was lower when audiences listened to music with complex auditory profiles (Bernardi et al., 2017), further suggesting physiological synchrony may be linked to structural coherence. ISC may additionally reflect familiarity and engagement: it is higher when listening to familiar musical styles, compared to unfamiliar styles; though upon repeated presentation, ISC drops with repetitions of familiar (but not unfamiliar) musical style (Madsen et al., 2019). However, many of these findings come from neural and laboratory contexts with mostly general descriptions of the stimuli. Thus, further research is required to deepen our understanding of music and ISC by investigating ISC with more portable methods such as physiology and exploring which musical features – characterised by a more comprehensive analysis – can evoke synchronised physiological responses in more naturalistic listening situations.

The overarching aim of the current study was to explore which musical feature(s) evoke systematic physiological responses during undisturbed, naturalistic music listening in a concert context. Participants attended one of three chamber music concerts, with live performances of string quintets by Beethoven (1770–1827), Dean (1961–), and Brahms (1833–1897) (four movements each), showcasing different musical styles (Viennese Classical, Contemporary, and Romantic, respectively) with varying tempo, tonality, compositional structure, character, and timbre. Psychological (emotion and absorption) ratings were collected from a short (2-minute) questionnaire immediately after each movement, and familiarity ratings were collected after each piece. Continuous physiological responses were measured throughout the concert, from which SCR, HR, RR, and EMG activity of 98 participants (Concert 1 [C1]: 36, C2: 41, C3: 21) was extracted. For each audience, ISC was calculated over a sliding window (5 musical bars long) for each physiological measure, representing the degree of collective synchrony of physiological responses over the time-course of each musical stimulus (see Fig. 1a). High and low physiological synchrony were operationalised based on criteria from Dmochowksi et al. (2012). Windows containing ISC values in the highest 20th percentile represented high synchrony (HS) windows, while windows with correlation ISC values in the 20th percentile centred around r = 0 (that is, lowest correlation values) represented low synchrony (LS) windows.

To characterise the music, quantitative values of low- and medium-level features as well as detailed descriptions of high-level, structural features were obtained using inter-disciplinary approaches of acoustic signal and score-based analyses, respectively. Acoustic features most commonly associated with orientation and arousal physiological responses (as described above) were extracted. Instantaneous tempo was calculated using inter- onset intervals (IOIs) between each beat to represent the speed. Root mean square (RMS) energy was computationally extracted to represent loudness (Laurier et al., 2009) of the music. As the centroid of the spectral distribution of an acoustic signal has been shown to be a main contributor to perception of timbre (Grey & Gordon, 1978) and timbral brightness (Alluri & Toiviainen, 2010; Wu et al., 2014), spectral centroid was computationally extracted (with higher values representing brighter timbre). Key clarity over time was calculated using an algorithm that correlates pitch class profiles (of the acoustic signal) with pre-defined key profiles (see Krumhansl, 2001; Krumhansl & Kessler, 1982; Toiviainen & Krumhansl, 2003). The correlation coefficient associated with the best-matched key represents a quantitative measure of clarity of key, which is highly (positively) correlated with perceptual ratings of key clarity (i.e., the reverse of harmonic ambiguity, Alluri et al., 2012). As we used naturalistic music, it was also important to consider stylistic and structural features of the music which cannot be analysed quantitatively (yet). Therefore, we prepared a detailed music theoretical analysis, indicating harmonic progressions, thematic and motivic relations, phrase rhythm, formal functions, and the overall repetition schema of the pieces (Hepokoski & Darcy, 2006; Zbikowski, 2002).

Based on the assumptions that ISC can identify systematic responses during naturalistic stimuli perception, and that certain physiological signatures and synchrony are related to musical features and self-reported states, the following hypotheses drove our research: (1) windows of high physiological synchrony – identified by ISC analysis – represent systematic physiological responses that are typically associated with event-related arousal responses; (2) highly correlated physiological responses during naturalistic music listening in a live concert context can be predicted by quantitative values of typically arousing acoustic features (higher RMS energy and spectral centroid, faster tempo, and lower key clarity), which may be modulated by the different styles. Additionally, we were (3) interested in the relations between ISC and higher-level musical features, such as compositional structure of the music. In light of the replicability crisis (Open Science Collaboration, 2015), we tested whether robust physiological responses to music would be consistent across repeated concert performances. Such an approach simultaneously provides data regarding the stability of using a concert as an experimental setting.

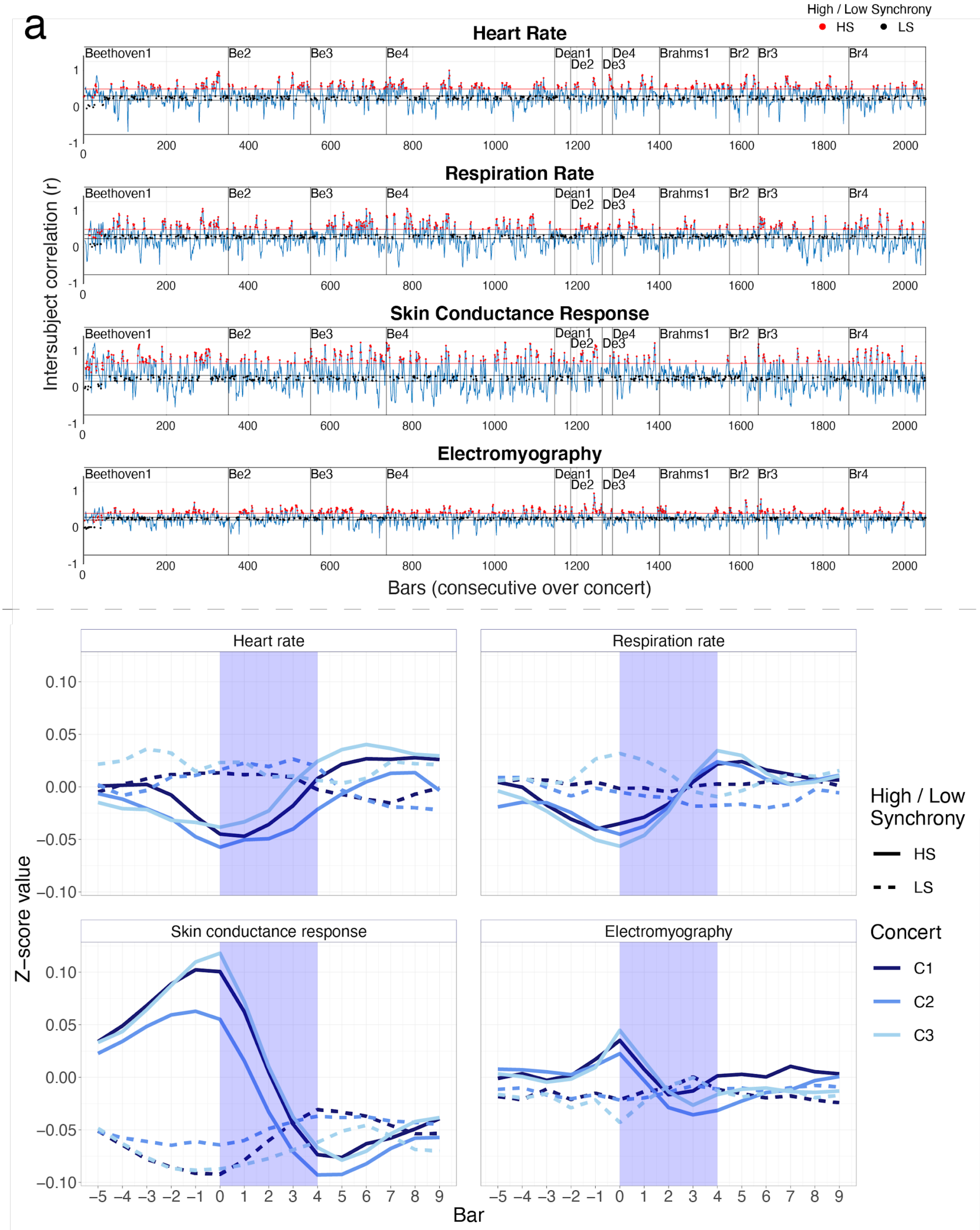


Figure 1. Inter-subject correlation (ISC) across concerts and bars of high- and low-synchrony. (**a**) ISC time courses for heart rate (HR, row 1), respiration rate (RR, row 2), skin conductance response (SCR, row 3), and electromyography activity of zygomaticus major ('smiling') muscle (EMG, row 4) for concert 1. Moments of high and low synchrony are marked with red and black dots, respectively. Red lines signify the 20th percentile threshold, while black lines signify the 20th percentile centred around $r = 0$. (**b**) Average high synchrony (HS) versus low synchrony (LS) of each physiological measure and concert. Moments of high and low synchrony are marked with solid and dotted lines,

respectively. Four musical bars precede (− 4 to 0) and follow (4–8) correlation windows (highlighted in blue box) with high ISC value starting from the first bar of correlation (bar0) to last bar of correlation window (bar4).

2 Results

Acoustic comparisons: differences between performances and styles. Before testing our hypotheses, we assessed whether the extracted acoustic features were comparable across performances. As the music was performed by professional musicians who were instructed to play as similarly as possible across the concerts, statistical tests confirmed our expectation that all concert performances would be acoustically similar. No significant differences occurred between performances for loudness, tempo, timbre, and length, with Pearson correlations of instantaneous tempo, timbre, and loudness between all performances reaching $r > 0.6$, $p < 0.001$ (see Supplementary Tables S1a, S1b, and S1c). This confirms that performances were comparable enough to allow for further statistical comparisons of listeners' physiology between audiences (i.e., that observed physiological responses are not attributable to unintended differences in the performances between concerts).

Because we hypothesised that style may play a role in driving physiological differences – and that certain styles are defined by differences in acoustic features – loudness, timbre, tempo, and key clarity were compared between styles. Contrasts revealed that Dean (Contemporary) had significantly lower RMS energy compared to Brahms (Romantic, $p < 0.05$). Dean also had significantly lower key clarity compared to both Beethoven (Classical, $p < 0.012$) and Brahms ($p < 0.001$), and significantly higher spectral centroid compared to Beethoven ($p < 0.037$) and Brahms ($p < 0.046$) (see Supplementary Figure S1 and Table S2a and S2b). Although tempo did not significantly differ between the styles (all $p > 0.327$), a division of tempi distribution in Beethoven and Brahms (Supplementary Figure S1) shows a typical composition practice of contrasting faster and slower movements in Classical and Romantic styles.

Summary. These acoustic checks confirm that our stimuli were comparable across concerts and that they offered a rich variation of acoustic features within and between pieces.

Physiological responses during moments of high versus low synchrony. Hypothesis 1 was assessed by comparing physiology in HS and LS windows. As shown in Fig. 1b, HR and RR were overall lower in HS compared to LS windows, confirmed by a significant main effect of Synchrony for HR in all concerts and for RR in C3. Significant Synchrony by Bar interactions and contrasts for RR in C2 and C3 suggested that breathing accelerates from the onset bar (bar0) to the last bar (bar4) of the HS window (see Tables 1, 2). An HR increase also seemed to occur (see Fig. 1b) but did not withstand Bonferroni correction.

SCR and EMG activity were overall higher in HS windows, compared to LS windows, confirmed by a significant main effect of synchrony for SCR and EMG in all three concerts. Significant Synchrony by Bar interactions and contrasts suggested that SCR (all concerts) and EMG activity (C2 and C3) decreased across HS windows (bar0–bar4, see Tables 1, 2). Looking at a wider range of 4 bars before and after the correlation window, it seems that ISC identifies the second half of an EMG peak amplitude and SCR peak, i.e., a momentary increase of sweat secretion (see Fig. 1b).

Summary. Compared to the LS moments, HS windows contained higher SCR and EMG activity, and increasing RR. Such responses correspond to a momentary activation of the sympathetic division of the ANS and have also been associated with self-reported arousal (Bradley, 2009; Bradley & Lang, 2000; Cacioppo et al., 2000; Hodges, 2010). These results support our first hypothesis that ISC can identify systematic, event-related physiological responses, indicative of increased arousal.

Acoustic properties as predictors of audience synchrony. Regarding our second hypothesis, quantitative values of tempo, key clarity, spectral centroid, and RMS during high and low synchrony windows were compared using logistic regression.

Single physiology measures. Tempo significantly predicted synchronised RR arousal responses in Beethoven and Brahms C1 and C2, and significantly predicted synchronised SCR arousal responses in Beethoven C2 and Brahms all concerts, where faster tempi increased the probability of synchronised RR and SCR responses across audience members (see Table 3 and Fig. 2). Slower tempi significantly increased probability of synchronised HR arousal responses in Beethoven, but only for C2 (see Table 3). Higher RMS energy significantly increased the probability of RR synchrony in Dean C1 and SCR synchrony in Beethoven C1 (see Table 3). No significant results occurred either for EMG synchrony or for spectral centroid or key clarity.

Multiple physiological measures. As HR, RR, and SCR, are all responses of the ANS, we assessed which musical features predicted a unified ANS response (i.e., when all three physiological measures were in synchrony simultaneously). Synchrony of the ANS as one entity (HR-RR-SCR) was not possible to model as no LS moments were found in the Dean piece for C1. Splitting ANS responses into paired combinations (HR-RR, HR-SCR, RR-SCR) yielded HS and LS moments in all three styles in all three concerts, allowing for further modelling. Logistic regressions revealed that faster passages (around 120 bpm, see Fig. 2) significantly increased probability of combined RR-SCR synchrony in Brahms C1 and C2 (see Table 4).

Summary. Although HR, RR, and SCR synchrony were predicted by tempo, RMS, and spectral centroid, the only result that remained consistent across at least two concerts was that synchrony of RR (in Beethoven and Brahms), SCR, and RR-SCR (in Brahms) were predicted by faster tempi. These findings partially support our second hypothesis, showing that one typically arousing music feature (faster tempo) increased the probability of two synchronised arousal-related responses (higher SCR amplitude and increasing RR) consistently, an effect which was modulated by style. Our hypothesis that louder RMS, brighter timbre and lower key clarity would predict physiological synchrony was not supported by the current results.

Concert	df	HR		RR		SCR		EMG	
		F	*p*	F	*p*	F	*p*	F	*p*
C1									
Synchrony	1	**13.316**	**<0.001**	5.146	0.023	**103.224**	**<0.001**	**20.044**	**<0.001**
Bar	4	**4.138**	**<0.001**	**5.485**	**<0.001**	**28.969**	**<0.001**	**5.335**	**<0.001**
Synch. × Bar	4	3.336	0.010	2.190	0.068	**26.392**	**<0.001**	**4.760**	**<0.001**
C2									
Synchrony	1	**24.426**	**<0.001**	8.045	0.005	**52.673**	**<0.001**	**12.815**	**<0.001**
Bar	4	1.743	0.137	**10.263**	**<0.001**	**26.981**	**<0.001**	**8.042**	**<0.001**
Synch. × Bar	4	0.846	0.495	**6.300**	**<0.001**	**19.372**	**<0.001**	**5.974**	**<0.001**
C3									
Synchrony	1	**10.065**	**0.002**	**25.527**	**>0.001**	**90.166**	**<0.001**	**29.240**	**<0.001**
Bar	4	3.739	0.005	**10.696**	**>0.001**	**24.780**	**<0.001**	**6.405**	**<0.001**
Synch. × Bar	4	**3.041**	**0.016**	**10.354**	**>0.001**	**15.977**	**<0.001**	**7.846**	**<0.001**

Table 1. ANOVA tests for linear models comparing physiology in 5 bar windows for Synchrony (HS/LS) across correlation windows in terms of Bar (0–4), calculated with the *Anova* function from the car package in R. Bonferroni-corrected threshold for significant effect was 0.05/12 = 0.004. Values highlighted in bold are statistically significant after Bonferroni corrections for multiple comparisons.

	HR					RR				
Pairwise comparison	**Estimate**	***SE***	***df***	***t***	***p***	**Estimate**	***SE***	***df***	***t***	***p***
C1										
HS–LS	**−0.036**	**0.007**	**4220**	**−5.073**	**<0.001**	−0.010	0.007	4000	−1.504	0.133
HS: bar0–bar4	−0.052	0.016	4220	−3.315	0.042	**−0.057**	**0.014**	**4000**	**−3.970**	**0.003**
LS: bar0–bar4	0.016	0.016	4220	0.997	1.00	−0.004	0.015	4000	−0.273	1.00
C2										
HS–LS	**−0.065**	**0.007**	**4225**	**−9.676**	**<0.001**	0.002	0.006	3705	−0.256	0.798
HS: bar0–bar4	−0.040	0.015	4225	−2.415	0.710	**−0.069**	**0.013**	**3705**	**−5.256**	**<0.001**
LS: bar0–bar4	−0.003	0.015	4225	−0.193	1.0000	0.013	0.015	3705	0.837	0.998
C3										
HS–LS	**−0.028**	**0.009**	**4195**	**−3.197**	**0.001**	**−0.028**	**0.008**	**3810**	**−3.518**	**<0.001**
HS: bar0–bar4	−0.062	0.012	4195	−3.217	0.060	**−0.091**	**0.017**	**3810**	**−5.495**	**<0.001**
LS: bar0–bar4	0.0170	0.019	4195	0.880	1.00	0.041	0.018	3810	2.248	0.424
	SCR					**EMG**				
Pairwise comparison	**Estimate**	***SE***	***df***	***t***	***p***	**Estimate**	***SE***	***df***	***t***	***p***
C1										
HS–LS	**0.071**	**0.009**	**4220**	**8.364**	**<0.001**	**0.014**	**0.006**	**4220**	**2.505**	**0.012**
HS: bar0–bar4	**0.174**	**0.019**	**4220**	**9.161**	**<0.001**	0.034	0.013	4220	2.683	0.33
LS: bar0–bar4	−0.062	0.019	4220	−3.252	0.052	−0.010	0.013	4220	−0.780	1.00
C2										
HS–LS	**0.025**	**0.007**	**4200**	**3.428**	**<0.001**	−0.001	0.006	4215	−0.084	0.933
HS: bar0–bar4	**0.148**	**0.017**	**4200**	**8.924**	**<0.001**	**0.054**	**0.012**	**4215**	**4.391**	**<0.001**
LS: bar0–bar4	−0.028	0.016	4200	−1.682	1.00	−0.010	0.012	4215	−0.845	1.00
C3										
HS–LS	**0.095**	**0.010**	**4205**	**9.796**	**<0.001**	**0.018**	**0.007**	**4215**	**2.504**	**0.012**
HS: bar0–bar4	**0.185**	**0.022**	**4205**	**8.530**	**<0.001**	**0.061**	**0.016**	**4215**	**3.758**	**0.008**
LS: bar0–bar4	−0.025	0.022	4205	−1.174	1.00	−0.030	0.016	4215	−1.868	1.00

Table 2. Pairwise comparisons of linear models comparing physiology in 5 bar windows for Synchrony (HS/ LS) across correlation windows in terms of Bar (0–4). HS–LS denotes the overall difference between HS and LS windows. Bar0–bar4 denotes the difference between the beginning and the end of the window separately in HS and LS windows. Contrasts (Bonferroni adjusted) were calculated with *emmeans* package in R. Values highlighted in bold are statistically significant after Bonferroni corrections for multiple comparisons.

	Beethoven			Dean			Brahms		
	Estimate	**SE**	***p* value**	**Estimate**	**SE**	***p* value**	**Estimate**	**SE**	***p* value**
Respiration rate									
C1									
(Intercept)	−1.211	0.897	0.177	1.059	2.150	0.6222	−0.728	1.255	0.562
Tempo	**0.0110**	**0.003**	**>0.001**	0.003	0.009	0.707	**0.018**	**0.005**	**>0.001**
RMS	−22.457	15.574	0.149	**212.796**	**50.117**	**>0.001**	−6.563	15.577	0.673
S. centroid	0.000	0.000	0.439	−0.000	0.006	0.559	−0.000	0.000	0.809
Key clarity	−0.151	1.114	0.891	−5.708	2.109	0.007	−1.218	1.569	0.438
C2									
(Intercept)	−0.343	1.034	0.740	2.376	1.898	0.211	−1.451	1.297	0.263
Tempo	**0.016**	**0.004**	**>0.001**	−0.006	0.009	0.502	**0.028**	**0.004**	**>0.001**
RMS	0.610	17.630	0.972	58.895	39.231	0.133	−17.878	17.983	0.320
S. centroid	−0.001	0.000	0.047	−0.001	0.001	0.049	−0.000	0.000	0.495
Key clarity	−0.371	1.111	0.738	−0.305	2.017	0.880	−1.089	1.713	0.525
C3									
(Intercept)	−1.614	0.953	0.090	−0.320	1.521	0.833	−0.3972	1.134	0.726
Tempo	0.007	0.003	0.016	0.004	0.007	0.537	0.008	0.003	0.0188
RMS	23.926	15.481	0.122	33.370	34.602	0.335	−22.432	15.333	0.143
S. centroid	0.000	0.000	0.618	−0.0002	0.001	0.628	−0.000	0.000	0.384
Key clarity	0.947	1.187	0.425	−0.478	1.685	0.777	0.929	1.415	0.511

	Beethoven			Dean			Brahms		
	Estimate	**SE**	***p* value**	**Estimate**	**SE**	***p* value**	**Estimate**	**SE**	***p* value**
Skin conductance									
C1									
(Intercept)	−1.959	0.913	0.032	3.522	1.813	0.052	−5.518	1.334	>0.001
Tempo	0.007	0.003	0.016	−0.015	0.008	0.0551	**0.028**	**0.005**	**>0.001**
RMS	**76.201**	**17.351**	**>0.001**	6.880	36.553	0.851	22.717	15.291	0.137
S. centroid	0.001	0.000	0.048	0.000	0.001	0.728	0.006	0.000	0.134
Key clarity	−0.626	1.024	0.541	−4.674	1.795	0.009	1.6770	1.532	0.274
C2									
(Intercept)	−1.015	0.913	0.266	−3.901	2.091	0.061	−5.567	1.470	>0.001
Tempo	**0.011**	**0.003**	**>0.001**	0.010	0.009	0.264	**0.036**	**0.006**	**>0.001**
RMS	17.223	15.542	0.268	29.043	43.571	0.505	44.689	16.671	0.007
S. centroid	−0.000	0.000	0.877	0.002	0.0001	0.006	0.000	0.000	0.438
Key clarity	−0.300	1.094	0.784	−1.475	2.094	0.481	0.302	1.740	0.862
C3									
(Intercept)	−0.660	0.883	0.455	0.512	1.576	0.745	−4.412	1.257	**>0.001**
Tempo	0.001	0.003	0.606	0.011	0.007	0.095	**0.015**	**0.004**	**>0.001**
RMS	21.770	15.067	0.148	−73.883	40.619	0.069	4.975	16.431	0.762
S. centroid	0.000	0.000	0.152	−0.001	0.001	0.174	0.001	0.000	0.209
Key clarity	−0.338	1.016	0.740	0.253	1.714	0.882	2.825	1.485	0.057

	Beethoven			Dean			Brahms		
	Estimate	SE	*p* value	Estimate	SE	*p* value	Estimate	SE	*p* value
Heart rate									
C1									
(Intercept)	−0.7443	0.916	0.416	5.459	2.028	0.007	0.996	1.180	0.399
Tempo	0.002	0.003	0.566	−0.008	0.007	0.261	−0.003	0.003	0.343
RMS	−25.496	16.908	0.132	−86.004	40.332	0.033	1.825	13.783	0.895
S. centroid	0.000	0.0003	0.103	−0.001	0.001	0.048	−0.001	0.000	0.161
Key clarity	0.008	0.979	0.993	−3.620	2.063	0.079	−0.036	1.532	0.981
C2									
(Intercept)	−0.435	0.899	0.629	0.480	1.490	0.747	0.367	1.213	0.762
Tempo	**−0.009**	**0.003**	**0.001**	−0.008	0.006	0.214	−0.003	0.003	0.406
RMS	20.089	14.853	0.176	25.795	32.573	0.428	−18.912	16.274	0.245
S. centroid	0.001	0.0003	0.009	0.0002	0.001	0.626	0.001	0.000	0.052
Key clarity	0.520	1.011	0.607	−1.765	1.376	0.200	−1.901	1.499	0.205
C3									
(Intercept)	2.544	0.907	0.005	1.229	1.726	0.477	−2.139	1.1802	0.070
Tempo	0.002	0.003	0.434	0.011	0.007	0.094	0.008	0.003	0.01
RMS	−47.392	15.412	0.002	−68.302	39.045	0.080	4.159	14.888	0.7800
S. centroid	−0.001	0.0002	0.029	0.001	0.001	0.307	0.000	0.000	0.761
Key clarity	−2.087	1.065	0.050	−0.798	1.947	0.682	1.457	1.433	0.309

Table 3. Logistic regressions for single physiological measure of respiration rate, skin conductance and heart rate synchrony per piece across all concerts (C1, C2, and C3). Synchrony (HS = 1, LS = 0) was the dependent variable, and tempo, key clarity, loudness, and spectral centroid (s. centroid) from the HS and LS bars as continuous predictors. The Bonferroni-corrected critical *p* value is 05/36 = 0.0014. Values highlighted in bold are statistically significant after Bonferroni corrections for multiple comparisons.

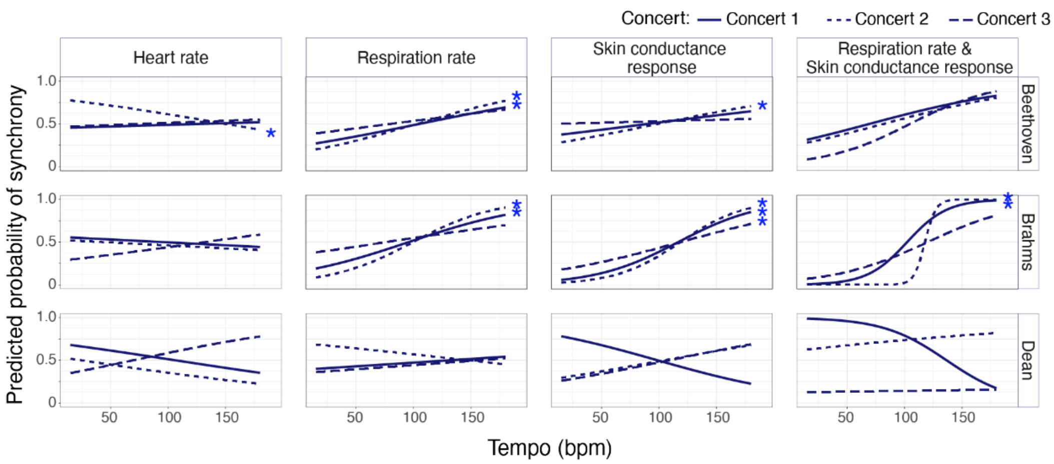


Figure 2. Logistic regression models with all music features (RMS, tempo, spectral centroid and key clarity) predicting high (1) versus low (0) synchrony across listeners, with probability curve of predictor tempo. Columns indicate

physiological measures of interest. Rows indicate each piece performed in each of the three concerts (indicated by line style). Here, we highlight the ability of tempo to predict synchrony. For full model results for all acoustic features, see Tables 3 and 4.

	Beethoven			**Dean**			**Brahms**		
	Estimate	**SE**	***p* value**	**Estimate**	**SE**	***p* value**	**Estimate**	**SE**	***p* value**
C1									
(Intercept)	−0.880	2.106	0.676	12.646	11.099	0.254	−6.255	4.129	0.130
Tempo	0.014	0.007	0.039	−0.036	0.049	0.461	**0.058**	**0.016**	**0.0004**
RMS	82.658	39.948	0.038	535.054	277.231	0.054	−7.109	36.676	0.8463
S. centroid	0.000	0.001	0.600	−0.001	0.002	0.508	0.001	0.002	0.323
Key clarity	−2.644	2.559	0.302	−18.948	9.346	0.043	−2.956	3.566	0.407
C2									
(Intercept)	0.935	2.486	0.707	−5.708	7.048	0.418	−13.149	7.235	0.069
Tempo	0.014	0.006	0.030	0.006	0.034	0.859	**0.211**	**0.067**	**0.002**
RMS	26.386	42.095	0.530	145.701	151.755	0.337	37.250	60.587	0.538
S. centroid	0.000	0.001	0.759	0.002	0.0028	0.549	−0.001	0.002	0.407
Key clarity	−4.533	3.077	0.141	3.612	6.376	0.571	−16.208	9.469	0.087
C3									
(Intercept)	−9.744	3.155	0.002	27.244	22.265	0.221	−4.988	3.006	0.097
Tempo	0.027	0.011	0.013	0.001	0.0212	0.943	0.025	0.011	0.031
RMS	36.114	37.380	0.334	−193.756	194.915	0.320	27.009	42.694	0.527
S. centroid	0.002	0.001	0.010	−0.0146	0.011	0.174	0.000	0.001	0.904
Key clarity	4.609	3.127	0.140	−11.609	12.505	0.353	2.445	2.863	0.393

Table 4. Logistic regressions for combined respiration rate and skin conductance response synchrony per piece across all concerts. Synchrony (HS = 1, LS = 0) was the dependent variable, and tempo, key clarity, loudness, and spectral centroid (s. centroid) from the HS and LS bars as continuous predictors. The Bonferroni-corrected critical p value is 0.05/27 = 0.0019. Values highlighted in bold are statistically significant after Bonferroni corrections for multiple comparisons.

Relationship between tempo and subjective experience. As we found a link between faster tempi and the highly synchronised SCR and RR increase (i.e., typical arousal responses) we wanted to further validate whether tempo was indeed related to self-reported emotion and engagement. Therefore, we correlated the mean tempo per movement with the psychological self-report data. Pearson correlations revealed that tempo positively correlated with engagement (e.g., feeling absorbed, concentrated), r=0.475, p=0.003, and positive high arousal emotions (e.g., energetic, joyful), $r = 0.513$, $p = 0.001$, with a medium and strong effect (Cohen, 1988), respectively.

Because the significance of our tempo results varied as a function of musical styles, we explored with descriptive statistics the possibility that familiarity with the piece might play a modulatory role. Beethoven and Brahms pieces were somewhat familiar to over a third of participants (41% and 35%, respectively), while the Dean piece was only somewhat familiar to 2.3% of participants (see Supplementary Table S3), suggesting that one interpretation why tempo significantly predicted physiological synchrony in the Beethoven and Brahms, but not Dean, could be due to differences in familiarity with the music.

Physiology synchrony across concerts in relation to the music theoretical analysis. With regard to our third hypothesis, we prepared a detailed music theoretical analysis of all the music (see Supplementary Table S4a, S4b, and S4c). We investigated musical events based on this analysis using moments of 'salient' physiological synchrony to find time points of interest in the music. Salient physiological responses were operationalised by two criteria: when (1) high physiological synchrony in any of the physiological measures occurred in all three concert audiences and (2) sustained synchrony occurred for more than one bar.

Overall, audience physiology seemed to synchronise around three types of musical events: (a) transitional passages with developing character, (b) clear boundaries between formal sections, and (c) phrase repetitions; all listed with descriptions in Supplementary Tables S5a, S5b, and S5c. Salient responses occurred during calming or arousing transitional passages (*calming*: Beethoven 1st movement, [Beethoven1], bars [b] 85–88; b287–293; Dean2, b70–71; Dean4, b75–77; Brahms4, b75–76; *arousing*: Beethoven1, b303–307; Dean2, b23–25; Brahms3, b6–8), characterised by a decrease (for calming passages) or an increase (for arousal passages) of loudness, texture, and pitch register. Other salient responses occurred when there was a clear boundary between functional sections, indicated through parameters such as a key change (e.g., between major and minor key in Beethoven3, b84–88; Brahms3, b58–61), a tempo change (e.g., Beethoven1 b328–331; Brahms4, b248–250), or a short silence (e.g., Beethoven1, b96–97; Beethoven4, b10–14). Lastly, salient responses occurred when a short phrase or motive was immediately repeated in a varied form, for example in an unexpected key, (e.g., Beethoven1, b35–37), elongated or truncated (e.g., Beethoven1, b85–88; 291–293), or with a different texture or pitch register (Brahms1, b90–91; Brahms3, b170–171). Since the immediate varied repetition of a short phrase is very common in Classical and Romantic styles, salient responses were also evoked when a phrase repetition occurred simultaneously with a transition or clear boundary (e.g., Beethoven1, b24–30; b136–138). With regard to style, the three categories are in line with the compositional conventions of the respective works: salient responses were found more often during transitions in the Romantic and Contemporary works, and during phrase repetitions and boundaries in the Classical work.

3 Discussion

This study assessed physiological aspects of the continuous music listening experience in a naturalistic environment using physiological ISC and inter-disciplinary stimulus analyses. We measured physiological responses of audiences listening to live instrumental music performances and examined which musical features evoked systematic physiological responses (i.e., synchronised responses, operationalised via ISC). Consistency of effects was assessed by repeating the same concert three times with different audiences. Importantly, no significant differences of length, loudness, tempo, or timbre across the concert performances were found, allowing us to assume that the musical stimuli were comparable across concerts.

Since previous research has identified typical physiological signatures as indices for felt arousal when listening to music, and since ISC is used to identify reliable responses across several individuals, we firstly hypothesised that ISC would identify similar physiological signatures of arousal. Our results supported this hypothesis: windows of high (compared to low) physiological synchrony contained significantly higher SCR/EMG and increasing RR, depicting similar physiological responses that have been previously related to self-reported arousal/tension evoked by music features like faster tempi (Chuen et al., 2016; Egermann et al., 2015; Gomez & Danuser, 2007; Khalfa et al., 2002; van der Zwaag et al., 2011), peak loudness (Grewe et al., 2007) as well as unexpected harmonic chords (Koelsch et al., 2008; Steinbeis et al., 2006) or notes (Egermann et al., 2013) in music. These patterns of physiological responses indicate that windows of high physiology ISC, therefore, likely correspond to event-related moments of increased arousal.

Our second hypothesis – that low- and medium-level acoustic features can predict high physiological synchrony – was partially supported. Specifically, logistic regressions revealed that one typically arousing musical feature – faster tempi – consistently (i.e., in at least two concerts) predicted synchronised RR and SCR arousal responses. Additionally, tempo consistently predicted a more general ANS response, that is, when both RR and SCR of audience members became synchronised simultaneously. In line with previous work showing that tempo and rhythm are the most important musical features in determining physiological responses (Gomez & Danuser, 2007), these findings suggest that tempo induces reliable physiological changes.

As faster tempo is typically perceived as more arousing (Coutinho & Cangelosi, 2011; Gabrielsson, 2012; Gabrielsson & Juslin, 1996; Gomez & Danuser, 2007), our result that faster music increased probability of high physiological synchrony supports previous research linking high ISC to increased arousal (Bracken et al., 2014; Dmochowski et al., 2012; Hasson et al., 2004). Our findings further support the idea that ISC is related to stimulus engagement (Madsen et al., 2019): as slower music increases mind-wandering (Taruffi et al., 2017), slower tempi may result in reduced attention to the music, leading to greater individual variability in physiological responses and subsequently lower ISC (Hasson et al., 2004; Simony et al., 2016). These connections between faster music with high physiological synchrony due to arousal and engagement were further confirmed by the psychological data, where faster tempi were positively correlated with audience members' self-reported engagement (e.g., high concentration and feelings of being absorbed) and high arousal positive emotions (e.g., feeling energetic or joyful). It is important to note that faster tempi (centred around 120 bpm/2 Hz in the current study) seem to be a more physiologically optimal range for entrainment to music (see Ding et al., 2017 for review). As entrainment is difficult at note rates under 1 Hz (at least for non-expert musicians Doelling & Poeppel, 2015), this may explain why synchrony did not occur in slower tempi of the current musical stimuli (which was centred around 50 bpm/0.83 Hz). It is therefore reasonable to speculate that entrainment, or the adaptation of autonomic physiological measures towards the musical tempo, might be a mechanism through which faster or optimally resonant tempi induce more-similar audience responses. In summary, we postulate that physiological synchrony may be predicted by faster music because of the physiological properties of the nervous system, which, when optimally driven, induce specific psychological experiences such as increased arousal and engagement.

It is of further interest that SCR and RR synchrony were more probable at faster tempi in the Classical and Romantic styles, but not in the Contemporary style, supporting previous studies where the same features evoke different physiological responses based on the style (Dillman Carpentier & Potter, 2007; Proverbio et al., 2015). As Beethoven and Brahms were rated as more familiar compared to Dean, ISC differences between styles in the current results support a recent study showing that ISC is linked to familiarity with musical style (Madsen et al., 2019). Additionally, Beethoven and Brahms have relatively stable meters (i.e., very few instantaneous tempo changes within movements), whereas many Dean passages contained unpredictable meters (e.g., the first movement has alternating bars of four, five, or six beats per bar) and frequent tempo changes within movements. In view of this, the fact that synchrony probability changes between styles could be due to stimulus coherence, corroborating studies showing that higher ISC occurs in more predictable contexts (Simony et al., 2016) and lower ISC occurs in versions of music where the beat is disrupted (Kaneshiro et al., 2020). However, since we presented only one work per style (and tempo changes were not evenly represented across these styles) – a compromise dictated by the constraints of a naturalistic concert setting, complimentary research using a wider range of pieces (in different styles) is required to further assess the effects of coherence and familiarity on synchrony.

Although we show that synchrony was predicted by tempo (depending on the style), the hypothesis that louder RMS, brighter timbre, and lower key clarity would predict physiological synchrony was not met. This was unexpected, as orienting/startle response research consistently shows that loudness evokes highly replicable physiological responses in a controlled tone sequence (Barry, 1975; Barry & Sokolov, 1993; Chuen et al., 2016). Because loudness in the current study was embedded in naturalistic music, our findings highlight the generalisability limitations of reductionist stimuli to real-world contexts (Nastase et al., 2020). Indeed, previous work has shown that environmental sounds and music evoke different physiological responses; an increase in HR (index of a startle response, Graham & Clifton, 1966) occurred with arousing noises (e.g., a ringing telephone or storm), but not with music (Gomez & Danuser, 2004). However, as we used such naturalistic music, it was important to not only explore quantitatively extracted low- and medium-level features, but also higher-level parameters in the music.

With regard to the hypothesis that synchrony corresponds to high-level/structural moments in the music, we observed that (from all moments in the music) transitional passages, clear boundaries, and immediate phrase repetitions in the music – identified using music theoretical analysis – coincided with highly synchronised physiological responses across concerts. Similar high-level features have been associated with physical responses such as shivers, laughter or tears (see Table IV in Sloboda, 1991) or specifically chills (Grewe et al., 2007), which also typically correspond with increased physiological arousal (Blood & Zatorre, 2001; Craig, 2005; Grewe et al., 2009; Guhn et al., 2007; Rickard, 2004; Salimpoor et al., 2009). However, none of the previous studies have derived these categories based on a detailed analysis of complete pieces and in the context of longer time spans.

In the current study, synchronised physiological responses occurred during arousing/calming transitional passages (characterised by changes in loudness, pitch register, and musical texture) and boundaries indicated by sudden tempo or key changes. Previous research has shown that unexpected musical events embedded in a predictable context may be perceived as arousing (Egermann et al., 2013; Koelsch et al., 2008; Steinbeis et al., 2006), further corroborating findings that high ISC occurs at arousing moments (Bracken et al., 2014; Dmochowski et al., 2012; Hasson et al., 2004). Our results additionally align with the notion that audience members collectively 'grip on' to loudness and texture changes (Phillips et al., 2020). The finding that physiological synchrony occurred during tempo changes, supports the fact that disruptions of temporal expectations affect ANS responses (Chuen et al., 2016; Fink et al., 2018) and EEG synchrony (Kaneshiro et al., 2020), where surprising events phase-reset ongoing physiological oscillations (for reviews, see Schroeder & Lakatos, 2009; Voloh & Womelsdorf, 2016), leading – at least briefly – to an increase in audience synchrony around moments of phase resetting.

We also found that momentary synchronised responses occurred during short phrases which were immediately repeated in a varied form, hinting at a general attention towards repetitions in music and a recognition of thematic connections over longer time spans (Margulis, 2014). Our analysis further suggests that an interplay of various musical features – in addition to the simple repetition – increase attention, and subsequently synchrony, of all audience members to these musical moments. For instance, high audience synchrony occurred in some of the structurally most important moments of Beethoven's 1st movement, where phrase repetition occurs simultaneously with a boundary (at the end of the exposition and with references to the primary theme at the end of the movement), and a boundary occurs simultaneously with transitional passages (deferred cadences; declined structural closures: e.g., b96–97, b301–302). The fact that high ISC occurred at phrase repetitions with an added novelty in the phrase (e.g., in a different key) – and also at important structural locations – not only supports research where novelty in stimuli evokes a physiological response (Barry, 1975; Barry & Sokolov, 1993; Bradley, 2009; Brown et al., 1991; Chuen et al., 2016; Dimberg, 1990; Frith & Allen, 1983; Graham & Clifton, 1966; Lyytinen et al., 1992; Sidle & Heron, 1976), but is also in line with music compositional practices, in which a composer tries to vary and develop thematic material (with different textures and/or harmonies) to keep listeners' interest.

In conclusion, by measuring continuous music listening experience in a naturalistic setting of a chamber music concert, we show that systematic synchronised physiological responses (corresponding to typical arousal responses) across audience members are predicted by tempo (depending on style) and are linked to structural transitions, boundaries, and phrase repetitions. Using naturalistic music in a concert environment is beneficial in that participants are likely to have more realistic and stronger responses. However, as this benefit makes our findings specific to the music we have used, future research should assess whether the current findings related to musical features and style are replicated with different (styles of) music. Additionally, further questions remain regarding the concert setting itself; for example, whether physiological effects and subjective experiences change with/without visual movements of the performer(s), with varying programming orders, and in different performance spaces (Tröndle et al., 2022; Wald-Fuhrmann et al., 2021). Exploring musical experiences from pre-recorded or live performances – with or without the co-presence of others – may prove an interesting future research direction, especially with regard to the COVID-19 pandemic and current transformations of the concert itself.

4 Method

Experimental procedures.

Experimental procedures are identical to Merrill et al. (2021). All experimental procedures were approved by the Ethics Council of the Max Planck Society and undertaken with written informed consent of each participant. All research was performed in accordance with the Declaration of Helsinki.

Participants.

129 participants attended one of three evening concerts. Some data were lost due to technical issues of server and user failures (*N* = 31) during data acquisition, leaving data from 98 participants for analysis. Gender and age were similarly distributed across concerts (see Table 5). Most participants reported that their highest level of education was a university degree, a German high school degree, or completed professional training. Musical Sophistication – assessed with the General Music Sophistication and Emotions subscales from the German version of the Goldsmiths Music Sophistication Index (Müllensiefen et al., 2014; Schaal et al., 2014) – was similarly moderate (see Appendix Table 3 in Merrill et al., 2021) across the three audience groups (see Table 5). Most participants reported that they regularly attend classical concerts and opera.

				Gold-MSI: emotion	Gold-MSI: general
Concert	Total *N*	Gender	Age	Mean score (SD)	Mean score (SD)
C1	36	F = 15, M = 17, *na* = 4	50% < 50 years old	33.53 (4.75)	69.84 (22.01)
C2	41	F = 16, M = 17, *na* = 8	50% < 55 years old	31.17 (6.85)	71.61 (21.97)
C3	21	F = 9, M = 12	50% < 40 years old	33.24 (5.45)	70.76 (19.33)

Table 5. Demographic information about participants in the current study, showing distribution of age, gender, and musical sophistication (general and emotions) across concerts. Participants were asked to report their age by selecting age group (within a 5-year range from 18 to 99, i.e., 18–22, 23–27, 28–32, etc.).

Concert context.

Three evening concerts (starting at 19.30 and ending at approximately 21.45) took place in a hybrid performance hall purpose-built for empirical investigations (the 'ArtLab' of the Max Planck Institute for Empirical Aesthetics in Frankfurt am Main, Germany). Care was taken to keep parameters (e.g., timing, lighting, temperature) as similar as possible across concerts. Professional musicians performed string quintets in the following order: Ludwig van Beethoven, op. 104 in C minor (1817), Brett Dean, 'Epitaphs' (2010), and Johannes Brahms, op. 111 in G major (1890), with a 20-minute interval between Dean and Brahms. This program was chosen to represent a typical chamber music concert.

Procedure

Participants were invited to arrive either one or one and half hour(s) before the concert began for physiological measurement preparation. Physiology was measured with a portable recording and amplifying device (https://plux.info/12-biosignalsplux) for the whole concert at 1000 Hz. Continuous blood volume pulse (BVP) was measured using a plethysmograph clip; respiration data were measured using a respiration belt (wrapped snugly around the lower rib cage); skin conductance was measured with electrodes placed on index and middle fingers of the non-dominant hand; and facial muscle activity was measured using electromyography, with adhesive electrodes placed above the zygomaticus major ('smiling') muscle on the left side of the face and ground placed on the mastoid. After each of the 12 movements, a short (2-min) pause was taken for the participants to fill in two questionnaires. The first questionnaire measured state absorption: eight items (e.g., 'I was completely absorbed by the music', 'My mind was wandering') were rated on a 5-point Likert scale from 'strongly disagree' to 'strongly agree' (Vroegh, 2018). The second questionnaire measured intensity of felt emotions using the **GE**neva **M**usic-**I**nduced **A**ffect **C**hecklist (GEMIAC, Coutinho & Scherer, 2017a), where intensity of 14 classes of feelings (e.g., energetic/ lively, tense/uneasy, nostalgic/sentimental) were rated on a 5-point Likert scale from 'not at all' to 'very much' (intensely experienced). Items were presented in German using comparable adjectives from the German version of the AESTHEMOS (Schindler et al., 2017). After each piece, participants rated how familiar the piece was ('Yes', 'No', and 'Not sure').

Data analysis.

Musical feature extraction.

Instantaneous tempo was calculated using inter-onset intervals (IOIs) between each beat (where beats were manually annotated by tapping each beat using Sonic Visualiser (Cannam et al., 2010). These IOIs were then converted to beats per minute (bpm). All other features were computationally extracted using the MIRToolbox (version 1.7.2) (Lartillot & Toiviainen, 2007) in MATLAB 2018b. In order to capture musical features meaningfully, different time windows were used to extract certain musical features. For features that change quickly on a timeframe of a less than a second – i.e., RMS energy (related to loudness), spectral centroid, brightness, and roughness (related to timbre) – a time window of 25 ms with a 50% overlap was used, as is typical in the music information retrieval literature (Alluri et al., 2012; Tzanetakis & Cook, 2002). Other features that are more context-dependent (that develop over a longer time frame), such as key clarity (i.e., the reverse of harmonic ambiguity) require a longer time-window, and were extracted similar to previous studies that assess the same musical feature, that is using a 3 s window (Alluri et al., 2012; Brattico et al., 2011) with 33% hop factor (overlap) (Mencke et al., 2019). As previous time-series analyses have parsed data into meaningful units of clause and sentence lengths (Kang & Wheatley, 2017), and as we wanted to aligned responses across concerts, each feature was averaged into a meaningful and comparable music unit: a bar (American: measure, on average 2 s long). It is worth noting that acoustic features can be distinguished between compositional features and performance features (Goodchild et al., 2019), where the former are represented in the musical score (such as harmony) while the latter include features that can change between performances, namely how loud and fast musicians may perform the music. Because key clarity is a compositional feature (i.e., does not change between performances), we had the same values across all three concert performances.

When checking for independence of features (Lange & Frieler, 2018), Pearson correlations revealed that RMS and roughness correlated highly as did brightness and spectral centroid (all $r > 0.7$, $p > 0.0001$) in all movements. As RMS and spectral centroid are features more commonly used, compared to roughness and brightness (Goodchild et al., 2019; Omigie et al., 2021), and spectral centroid seems to best represent timbre (Grey & Gordon, 1978; McAdams et al., 1995) and brightness (Alluri & Toiviainen, 2010) perception, we kept only key clarity, RMS, spectral centroid, and tempo for further analysis. To check performance feature similarity between concerts, features were compared with concert (C1, C2, C3) as the independent variable. Pearson correlations were used to assess similarity of acoustic features over time between concerts. Correlations were considered adequate if they met a large effect size of concert $r > 0.5$ (Cohen, 1988). To compare acoustic features per style, linear mixed models with fixed effect of the works (Beethoven, Brahms, and Dean) and random intercept of movement were constructed with each acoustic feature (as the dependent variables) per concert.

Physiology pre-processing.

Physiological data were pre-processed and analysed in MATLAB 2018b. Data were cut per movement. Missing data (gaps of less than 50 ms) were interpolated at the original sampling rate. Fieldtrip (Oostenveld et al., 2011) was used to pre-process BVP, respiration, and EMG data. BVP data were band-pass filtered between 0.8 and 20 Hz (4th order, Butterworth) and demeaned per movement. Adjacent systolic peaks were detected to obtain inter-beat intervals (IBIs) and an additional filter was added to remove any IBIs that were shorter than 300 ms, longer than 2 s, or had a change of more than 20% between adjacent IBIs (typical features of incorrectly identified IBIs, Piskorski & Guzik, 2005). After visual inspection and artefact removal, IBIs were converted to continuous heart rate (HR) by interpolation. Respiration data were low pass filtered (0.6 Hz, 6th order, Butterworth) and demeaned. Maximum peaks were located, and respiration rate (RR) was inferred by the peak intervals. EMG activity was band-pass filtered (between 90 and 130 Hz, 4th order, Butterworth), demeaned, and the absolute value of the Hilbert transform of the filtered signal was extracted and smoothed. Skin conductance data were pre-processed using Ledalab (Benedek & Kaernbach, 2010) and decomposed into phasic and tonic activity. As we were interested in event-related responses, only (phasic) skin conductance responses (SCR) were used in further analyses. All pre-processed physiological data (SCR, HR, RR, EMG) were resampled at 20 Hz (Bradley & Lang, 2000), z-scored within participant and movement, and averaged into bins per bar.

ISC analysis.

We calculated a time-series ISC based on Simony et al. (2016) by forming $p \times n$ matrices (one for each SCR, HR, RR, and EMG, and for each of the twelve movements per concert), where p is the physiological response for each participant over n time points (bars across the movements). Correlations were calculated over a sliding window 5-bars long (approximately 10 s; the average bar length across the whole concert was 2 s), shifting one bar at a time. Fisher's r-to-z transformation was applied to correlation coefficients per subject, then averaged z values were inverse transformed back to r values. The first 5 bars and the last 5 bars of each movement were discarded to remove common physiological responses evoked by the onset/offset of music (Wilson et al., 2008). ISC values per movement were concatenated within each concert (2238 bars), giving four physiological ISC measures per concert. These

ISC traces represent the similarity of the audience members' physiological responses over time (see Fig. 1a).

Following Dmochowski et al. (2012), high and low synchrony were defined using 20th percentiles. Windows containing the highest 20th percentile of ISC values were categorised as high synchrony (HS) windows. Windows with values in the lowest 20th percentile of correlation values (i.e., ISC values within a 20th percentile where *r* was centred around zero) were categorised as low synchrony (LS) windows. To obtain instances of overall ANS synchrony (i.e., across multiple physiological measures simultaneously), we identified where HS/LS moments of one physiological measure coincided with another physiological measure. Physiological responses at points of HS and LS were compared using linear models with fixed effects Synchrony (high/low) and Bar (bars 0–4). To investigate whether acoustic features predicted synchrony (HS/LS) of physiological responses across audience members, tempo, RMS energy, key clarity, and spectral centroid in bars of HS/LS were recovered. By dummy-coding Synchrony as a binary variable (HS as 1; LS as 0) logistic regression models were constructed to predict Synchrony for each physiological measure (dependent variable) with continuous predictors of tempo, key clarity, loudness, and spectral centroid from the HS and LS bars (all features were included, as perceived expression in music tends to be determined by multiple musical features (Gabrielsson, 2012). (N.B.: no random intercept of movement was included, because not all movements contained HS/LS epochs). As we expected style to modulate the effect of these acoustic features in predicting synchrony, models were run separately per piece and concert.

Subjective ratings.

A factor analysis was conducted for all ratings from the absorption and GEMIAC scales. Five latent variables were identified: (1) Positive high arousal emotions (energetic, powerful, inspired, joyful, filled with wonder, enchanted); (2) Negative high arousal emotions (tense, agitated, negative loadings of liking and feeling relaxed); (3) Mixed valence, low arousal emotions (relaxed, nostalgic, melancholic, feelings of tenderness, moved); (4) Engagement (concentrated, forgetting, absorbed, liking, with negative loadings of bored, indifferent, and mind-wandering); and (5) Dissociation (forgetting being in a concert and forgetting surroundings); see (Merrill et al., 2021) for their exact loadings. We (Pearson) correlated participants' factor scores with musical features.

Music theoretical analysis.

The scores of all works were analysed according to widely used methods for the respective styles. Musical events were analysed on the beat level (harmonic changes, cadences, texture changes, motivic relation (Schönberg, 1999) and grouped into larger sections (thematic relations, formal functions/action spaces, repetition schemata (Hepokoski & Darcy, 2006; Zbikowski, 2002). The performance recordings served as references for passages which could have been interpreted equivocally in the score. After the analysis, passages involving high physiological synchrony were marked. These musical features were compared with each other, categorized across styles, and finally reduced to three categories: (a) transitional passages, (b) clear boundaries between formal sections, and c) phrase repetitions.

Statistical analyses.

Statistical analyses were conducted in R (R Development Core Team, 2020). Pearson correlations were computed using *corr.test* in the *psych* package (Revelle, 2021) and adjusted for false discovery rate using the Benjamin-Hochberg procedure. Linear (mixed) models were constructed using the *lme4* package (Bates et al., 2015); *p* values were calculated with the *lmerTest* package (Kuznetsova et al., 2017) and using the *Anova* function in the *car* package (Fox & Weisberg, 2018). Contrasts were assessed with the emmeans function (*emmeans* package, Lenth, 2021). Logistic regression models were run using a general linear model with a logit link function. Significance thresholds for *p* values for ANOVA, contrasts, and logistic regression models were adjusted using Bonferroni corrections.

5 Supplementary Materials

Supplementary Table S1a. ANOVA for performance (acoustic) differences across concerts. ANOVA was calculated using *Anova* function in car package in R.

		Beethoven		Brahms		Dean	
	df	F	p	F	p	F	p
RMS	2	2.14	0.12	0.23	0.79	0.12	0.89
Spectral centroid	2	1.78	0.17	1.54	0.21	2.82	0.06
Tempo	2	2.56	0.08	0.02	0.98	2.31	0.10

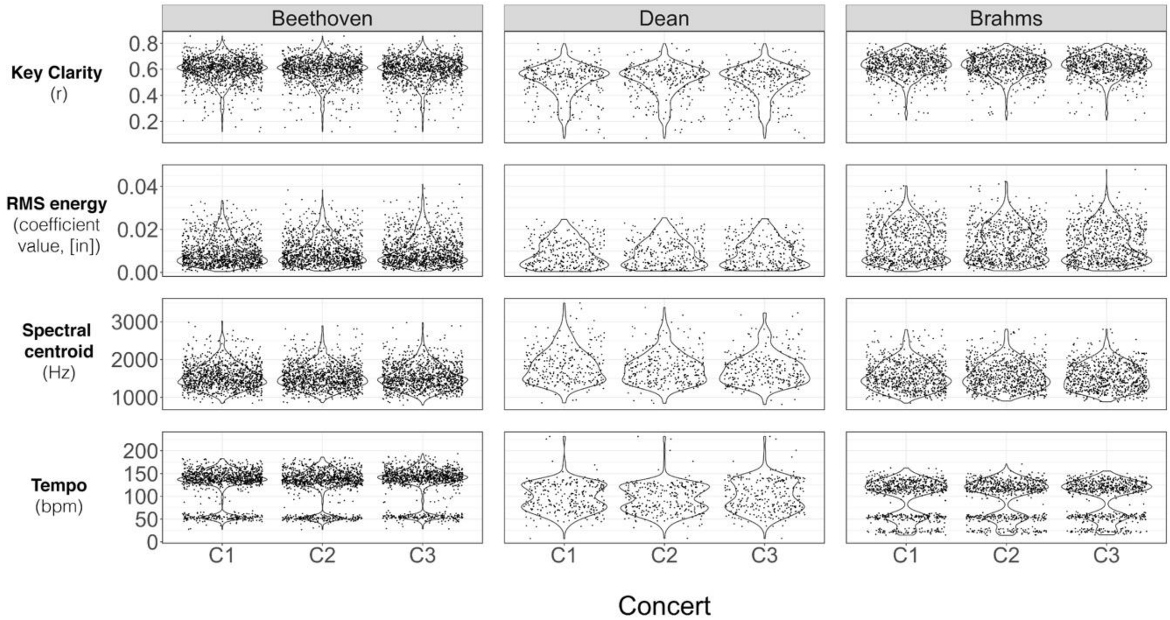
Supplementary Table S1b. Lengths of separate movements of works.

	Beethoven movements				Dean movements				Brahms movements			
	1st	2nd	3rd	4th	1st	2nd	3rd	4th	1st	2nd	3rd	4th
C1	7.80	8.15	3.76	6.37	2.53	3.87	3.90	3.39	10.17	7.15	5.37	4.94
C2	7.95	8.36	3.74	6.50	2.70	3.87	3.93	3.49	10.24	7.07	5.39	4.95
C3	7.75	8.02	3.71	6.23	2.68	3.87	3.79	3.14	10.03	6.97	5.46	4.93

Supplementary Table S1c. Correlations between features within movements. Bonferroni corrected *p* value was *p* = .0003. All correlations significance values were still under this threshold and, therefore, significant. Correlations were run using *corr.test* function from the *psych* package in R.

	RMS			Spectral Centroid			Tempo		
	Comparison	*r*	*p*	Comparison	*r*	*p*	Comparison	*r*	*p*
Bee1	C1 - C2	.947	> .0001	C1 - C2	.8760	> .0001	C1 - C2	.822	> .0001
	C1 - C3	.946	> .0001	C1 - C3	.889	> .0001	C1 - C3	.830	> .0001
	C2 - C3	.957	> .0001	C2 - C3	.927	> .0001	C2 - C3	.841	> .0001
Bee2	C1 - C2	.875	> .0001	C1 - C2	.901	> .0001	C1 - C2	.858	> .0001
	C1 - C3	.855	> .0001	C1 - C3	.908	> .0001	C1 - C3	.861	> .0001
	C2 - C3	.839	> .0001	C2 - C3	.900	> .0001	C2 - C3	.832	> .0001
Bee3	C1 - C2	.887	> .0001	C1 - C2	.940	> .0001	C1 - C2	.811	> .0001
	C1 - C3	.884	> .0001	C1 - C3	.910	> .0001	C1 - C3	.821	> .0001
	C2 - C3	.916	> .0001	C2 - C3	.882	> .0001	C2 - C3	.824	> .0001
Bee4	C1 - C2	.922	> .0001	C1 - C2	.920	> .0001	C1 - C2	.684	> .0001
	C1 - C3	.904	> .0001	C1 - C3	.906	> .0001	C1 - C3	.628	> .0001
	C2 - C3	.902	> .0001	C2 - C3	.888	> .0001	C2 - C3	.622	> .0001
Bra1	C1 - C2	.964	> .0001	C1 - C2	.978	> .0001	C1 - C2	.848	> .0001
	C1 - C3	.947	> .0001	C1 - C3	.962	> .0001	C1 - C3	.844	> .0001
	C2 - C3	.956	> .0001	C2 - C3	.973	> .0001	C2 - C3	.881	> .0001
Bra2	C1 - C2	.975	> .0001	C1 - C2	.965	> .0001	C1 - C2	.893	> .0001
	C1 - C3	.974	> .0001	C1 - C3	.961	> .0001	C1 - C3	.871	> .0001
	C2 - C3	.984	> .0001	C2 - C3	.977	> .0001	C2 - C3	.932	> .0001
Bra3	C1 - C2	.957	> .0001	C1 - C2	.933	> .0001	C1 - C2	.595	> .0001
	C1 - C3	.940	> .0001	C1 - C3	.939	> .0001	C1 - C3	.692	> .0001
	C2 - C3	.950	> .0001	C2 - C3	.932	> .0001	C2 - C3	.645	> .0001
Bra4	C1 - C2	.957	> .0001	C1 - C2	.936	> .0001	C1 - C2	.843	> .0001
	C1 - C3	.961	> .0001	C1 - C3	.920	> .0001	C1 - C3	.831	> .0001
	C2 - C3	.952	> .0001	C2 - C3	.931	> .0001	C2 - C3	.848	> .0001
Dea1	C1 - C2	.848	> .0001	C1 - C2	.751	> .0001	C1 - C2	.987	> .0001
	C1 - C3	.876	> .0001	C1 - C3	.843	> .0001	C1 - C3	.992	> .0001
	C2 - C3	.939	> .0001	C2 - C3	.895	> .0001	C2 - C3	.991	> .0001
Dea2	C1 - C2	.999	> .0001	C1 - C2	.982	> .0001	C1 - C2	1	> .0001
	C1 - C3	.998	> .0001	C1 - C3	.976	> .0001	C1 - C3	1	> .0001
	C2 - C3	.998	> .0001	C2 - C3	.990	> .0001	C2 - C3	1	> .0001
Dea3	C1 - C2	.982	> .0001	C1 - C2	.972	> .0001	C1 - C2	.940	> .0001
	C1 - C3	.971	> .0001	C1 - C3	.955	> .0001	C1 - C3	.858	> .0001
	C2 - C3	.975	> .0001	C2 - C3	.944	> .0001	C2 - C3	.913	> .0001
Dea4	C1 - C2	.947	> .0001	C1 - C2	.936	> .0001	C1 - C2	.936	> .0001
	C1 - C3	.944	> .0001	C1 - C3	.931	> .0001	C1 - C3	.891	> .0001
	C2 - C3	.940	> .0001	C2 - C3	.936	> .0001	C2 - C3	.887	> .0001

Note. Dean 2 was played via a recording, hence the perfect correlation for tempo across concerts (and near perfect for loudness and spectral centroid, with slight possible deviations in the audio recording).



Supplementary Figure S1. Acoustic features per bar, per piece, per concert. Top to bottom panels show Key clarity, RMS energy, Spectral centroid, and tempo values per bar. Left panels show values for Ludwig van Beethoven (String Quintet in C minor, op. 104, 1817), middle panels for Brett Dean (Epitaphs, 2010), and right panels for Johannes Brahms (String Quintet in G major, op. 111, 1890). Separate violin plots show different concerts.

Supplementary Table S2a. Musical feature differences between styles. ANOVA was calculated using *Anova* function in *car* package in R.

		C1		C2		C3	
Acoustic feature	df	χ^2	*p*	χ^2	*p*	χ^2	*p*
RMS	**2**	**9.832**	**.007**	**9.853**	**.007**	**8.788**	**.012**
Spectral centroid	**2**	**21.630**	**< .001**	**12.165**	**.002**	**13.736**	**.001**
Key clarity	**2**	**28.738**	**> .001**	**28.738**	**> .001**	**28.738**	**> .001**
Tempo	2	2.136	.343	2.125	.346	2.237	.327

Note. As key clarity was the same across all concerts, there was no change in the difference between pieces between concerts.

Supplementary Table S2b. Pairwise comparisons of models comparing acoustic features between different pieces. Contrasts (Bonferroni adjusted) were calculated with *emmeans* package in R.

Feature	Concert	Pairwise comparison	Estimate	SE	df	*t*	*p*
RMS	C1	Beethoven-Brahms	-0.004	0.002	8.46	-2.423	.120
		Beethoven– Dean	0.001	0.002	9.22	0.571	1.00
		Brahms – Dean	**0.005**	**0.002**	**9.42**	**2.923**	**.049**
	C2	Beethoven-Brahms	-0.004	0.002	8.40	-2.361	.101
		Beethoven– Dean	0.001	0.002	9.25	0.676	1.00
		Brahms – Dean	**0.005**	**0.002**	**9.48**	**2.959**	**.045**
	C3	Beethoven-Brahms	-0.003	0.002	8.40	-2.090	.205
		Beethoven– Dean	0.001	0.002	9.25	0.836	1.00
		Brahms – Dean	**0.005**	**0.002**	**9.48**	**2.856**	**.054**
Spectral centroid	C1	Beethoven-Brahms	0.54	76.6	8.39	0.007	1.00
		Beethoven– Dean	**-320.36**	**78.7**	**9.26**	**-4.072**	**.008**
		Brahms – Dean	**-320.90**	**79.2**	**9.49**	**-4.053**	**.008**
	C2	Beethoven-Brahms	-12.2	87.3	8.54	-0.139	1.00
		Beethoven– Dean	**-277.0**	**89.0**	**9.18**	**-3.112**	**.037**
		Brahms – Dean	**-264.**	**89.4**	**9.35**	**-2.962**	**.046**
	C3	Beethoven-Brahms	-0.828	75.8	8.42	-0.011	1.00
		Beethoven– Dean	**-252.462**	**77.7**	**9.24**	**-3.251**	**.029**
		Brahms – Dean	**-251.633**	**78.1**	**9.45**	**-3.221**	**.030**
Key clarity		Beethoven-Brahms	-0.031	0.021	8.40	-1.495	.541
		Beethoven– Dean	**0.080**	**0.021**	**9.25**	**3.791**	**.012**
		Brahms – Dean	**0.111**	**0.021**	**9.48**	**5.216**	**.001**

Note. As key clarity was the same across all concerts, there was no change in the difference between pieces between concerts.

Supplementary Table 3. Familiarity ratings across concerts that were collected after each piece to the question if participants knew the piece (with three possible choices: yes, no, not sure).

	Familiar	Not sure if familiar	Unfamiliar
Beethoven	12.5%	28.4%	59.1%
Brahms	15.9%	19.3%	65.9%
Dean	0%	2.3%	97.7%

Supplementary Table S4

Ludwig van Beethoven: Quintet op. 104, c minor. Music theoretical analysis of Mvt. 1, Allegro con brio, bars 1-137 (exposition)

Table 4a: Complete movement, division of formal sections

From bar	To bar	Form part	Length in bars
1	30	Exposition - Primary zone	*30*
31	58	Exposition - Transition	*28*
59	124	Exposition - Secondary zone	*66*
124	137	Exposition - Closing zone	*14*
138	213	Development zone	*76*
214	223	Recapitulation - Primary Zone	*10*
224	261	Recapitulation - Transition	*38*
262	327	Recapitulation - Secondary Zone	*66*
328	360	Recapitulation - Closing Zone	*33*

Table 4b: Sample analysis of the exposition with broad function, harmonic context / development, phrase structure and dynamic relations

From bar	To bar	Event bar.beat	Function	Phrase structure	Harmonic context	Dynamics (descriptive)
1	30		Exposition - Primary zone		i = c minor	
1	10		P^{0}	4+2+2+2	c minor	soft
		3-4	basic idea: arpeggiated V (sixth chord)			
		5-6	b.i. repetition, on VI (sixth chord)		A♭ major	
		9.1	delay, ornamented cadence			soft
		9.2	ornament: highpoint of the melody, seventh			
		9.3	rethoric delay			
		10.1	HC, 9-8 suspension, fermata / delay		G major	soft
11	14		$P^{pres.\ 1}$: compound basic idea	2+2	c minor	soft
		13-14	HC, 4-3 suspension			
15	18		$P^{pres.\ 2}$: prolinging the tonic	2+2		soft
		17-18	IAC, 4-3 suspension			soft
19	30		$P^{cont.}$, pedal point on dominant, pendulum V-i (D-t), pre-cadence	6+2+2+2		increasing
		19.1	start of pedal point			accent
		21	increasing density, syncopations on 19.2, 20.2, 21.2 etc.			increasing
		27.1	dynamic maximum			strong
		27.3	dynamic maximum			strong
		28.1	dynamic maximum			strong
		29-30	HC		G major	strong
31	58		Exposition - Transition		i --> III	
31	39		TR-P^{0}, transposed, shortened	4+2+2	A♭ major	soft
		31.1	harmonicaly surprising shift to VI (tG)			strong
		33-34	compare bars 3-4, motive: arpeggiated V (sixth chord)			
		35-36	compare bars 5-6, motive repetition, on VI (sixth chord)		F♭ major	
		37.1	ante penultima - compare to bar 7			strong
		38.1	penultima, leadington in the bass			strong
37	39		chromatic bassline progression to B♭, dominant of III (D)tP	2+1		increasing
39	57		"standing on the dominant", structural pedal point on B♭		B♭ major	
40	43		TR-$P^{1\ ant}$	2+2	(V-i)	increasing
44	47		TR-$P^{1\ cons}$	4+4		increasing
47	53		TR-precadence, Takterstickung	4+3	(V-i)	strong
		53.1	MC declined			strong
53	56		cadence, falling bassline	4+2		decreasing
57	58		Takterstickung, MC-fill			
		57.1	MC deformed: HC V of III (tP)		B♭ major	
59	124		Exposition - Secondary zone		I = E♭ major	
59	66		$S^{1.1}$, antecedent+consequent	4+4	E♭ major	soft
67	75		$S^{1.2}$, transposed repetition of S^{1}, modulation to IV	4+4+1		soft
		69.1	beginning: new tonal region: IV		A♭ major	soft
		75.1	IV: IAC			
		75	sustaining transitional bar			
76	79		S^{2ant}	2+2		decreasing
		79.1	IV: HC			
80	91		S^{2con}, pedalpoint on E♭, modulation to V dimin (1-->7th)	2+2+2+2+1+1+1+1	e♭ - a♭ - Fdim.	decreasing
		91.1	accent, syncopated phrase rhythm			strong
92	97		pre-cadence V, harmonic rhythm accelerated	4+2	B♭ major	strong
		96-97	V: HC, 64-53 suspension			
		98.1	I: evaded PAC, declined EEC		E♭ major	strong
98	109		sequences of P^{0}, pre-cadential	4+2+2+1+1+1+1		strong
110	118		$S^{3.1}$, augmentation of ornamenting neighbournotes, IAC	4+4		
		111.1	Neopolitan chord		F♭ major	
118	124		$S^{3.2}$, augmentation of ornamenting neighbournotes, IAC	3+3		arching
		118.1	Deceptive cadence		C♭ major	
124	137		Exposition - Closing zone	4+4	I = E♭ major	increasing
		124.1	PAC, EEC		E♭ major	strong
132	137		cadence	2+2+2		strong

Table 4c: Terminology & References

Material:	P, S, TR = Primary theme, Secondary theme, Transition (Hepokoski 2006)
Temporal labels:	$P^{0,\ 1,\ etc.}$ = Primary Theme and its consecutive sections (0 as introduction, 1 as first part of theme one, etc.), e.g. TR-P^{0} = section of transition with material that originates in the primary zone (Hepokoski 2006)
Phrase structure:	ant. / cons. / pres. / cont. = antecedent / consequent / presentation / continuation. (Caplin 1998)
Cadences:	HC = half cadence, IAC = imperfect authentic cadence, PAC = perfect authentic cadence, DC = deceptive cadence
Functional cadences:	EEC = essential expositional closure, ESC = essential structural closure, MC = medial caesura (Hepokoski 2006)

Caplin, W. E. (1998). *Classical Form : A Theory of Formal Functions for the Instrumental Music of Haydn, Mozart, and Beethoven.* New York: Oxford University Press

Hepokoski, J. A., & Darcy, W. (2006). *Elements of sonata theory: Norms, Types, and Deformations in the Late Eighteenth-Century Sonata* . New York: Oxford University Press.

Supplementary Table S5a. Musical descriptions of bars with salient physiological responses: Beethoven

Ludwig van Beethoven: String Quintet in C minor op. 104 (1817)			
Movement	**Bars**	**Categories**	**Description**
1st *	24-30	Phrase repetition Transition	Exposition: Increasing texture, dynamics & harmonic rhythm, modulation and half-cadence bar30 ending the primary-theme zone (P)
	35-37	Phrase repetition	Exposition: Chromatic sequence at opening of transition (T)
	85-88	Phrase repetition Transition	Exposition: Decreasing texture & dynamics, bar-wise repetition of motif, elongation & prolongation of secondary diminished dominant
	96-97	Boundary	Exposition: Strong half-cadence, essential expositional closure (EEC) declined in b97
	136-138	Phrase repetition Boundary	Exposition: Increasing texture & dynamics, syncopations, cadence ending the closing zone (C) of exposition. First chord of development section.
	287-290	Phrase repetition Transition	Recapitulation: Decreasing texture & dynamics, bar-wise repetition of motif, prolongation of secondary diminished dominant
	291-293	Phrase repetition	Recapitulation: Elongation & prolongation of secondary diminished dominant
	301-302	Boundary	Recapitulation: essential structural closure (ESC) declined in b300, beginning of strong transitional section, motivic development of P in unison
	303-307	Transition	Recapitulation: strong transitional section, motivic development of P in unison
	328-331	Boundary	Coda: Reference to opening bars with fermata & tempo change (150bpm to Adagio, ca. 80 bpm and back) after final cadence in b327
2nd	62-64	Boundary	Variation 1: Perfect authentic cadence (PAC), end of var 1
	126-130	Transition Boundary	Variation 3: increase of dynamic (*morendo*), PAC b129 ending var. 3 in minor. Variation 4: Key, texture, tempo change, beginning of var.4 in major
3rd	65-67**	Phrase repetition	Menuetto 1, B-part 2nd time: standing on the dominant after half cadence (HC) b64, reference to beginning of B-part b55-58
	84-88	Boundary	Menuetto 1, A'-part 2nd time: end of menuetto (minor), beginning of Trio (major). Change of key, register, texture, phrase rhythm
4th *	10-14	Boundary Transition	Exposition: beginning of primary theme one (P1) after intro & general rest in b8

	150-153	Phrase repetition Boundary	Development: continuation within P0 at the beginning of development, half cadence, general rest
	187-188	Phrase repetition Boundary	Development: rotation of subordinate theme (S), second pair of of six-bar chromatic sequence of S

**Note*: no repeat. ***Note*: Bars in relation to repeats

Supplementary Table S5b. Musical descriptions of bars with salient physiological responses: Dean

Brett Dean: Epitaphs (2010)			
Movement	**Bars**	**Categories**	**Description**
2nd	23-25	Transition	Increasing texture density, rhythmically clearer than previous context
	61-62	Transition Boundary	Closing phrase with decreasing loudness cadence and *finalis* with fermata
	70-71	Transition	Increasing loudness and dissonance, varied repetition, change of pitch class quality & timbre (*increasingly raw*)
4th	75-77	Boundary Transition	Sudden decrease of dynamics, change of texture & timbre Glissando over 3+3+2 pattern

Supplementary Table S5c. Musical descriptions of bars with salient physiological responses: Brahms

Johannes Brahms: String Quintet in G major op. 111 (1890)			
Movement	**Bars**	**Categories**	**Description**
1st *	90-91	Phrase repetition	Development: combined P/S space, homophonic texture
3rd	6-8	Transition	1st minor part, A: modulation to half cadence in the middle of 8-bar continuation. Small increase of dynamics, descending line in upper melody
	58-61**	Boundary	1st minor part, B: end of first, beginning of second repeat. Final cadence with *finalis* b58 lowest pitch of the melody, change to major b59, change of register and texture b61
	170-171**	Boundary Phrase repetition	2nd minor part, A: end of 12-bar antecedent on HC, beginning of consequent (repeat of antecedent), change of register to upper octave
	218-219	Boundary	2nd major part, coda: end of 2nd minor part, beginning of coda, change of key

4th	75-76	Transition	Exposition: within coda, change of texture & timbre, decrease of dynamics, modal cadence with 5-6 progression
	248-250	Boundary	Coda: beginning of *stretta* (120 bpm to 140 bpm), homophonic, clear rhythm

**Note*: no repeat. ***Note*: Bars in relation to repeats

Chapter 4

Aesthetic and physiological effects of naturalistic multimodal music listening

Based on:

Czepiel, A., Fink, L. K., Seibert, C., Scharinger, M., & Kotz, S. A. (2023). Aesthetic and physiological effects of naturalistic multimodal music listening. *Cognition*, *239*, 105537. https://doi.org/10.1016/j.cognition.2023.105537

Abstract

Compared to audio-only (AO) conditions, audio-visual (AV) information can enhance the aesthetic experience of a music performance. However, such beneficial multimodal effects have yet to be studied in naturalistic music performance settings. Further, peripheral physiological correlates of aesthetic experiences are not well-understood. Here, participants were invited to a concert hall for piano performances of Bach, Messiaen, and Beethoven, which were presented in two conditions: AV and AO. They rated their aesthetic experience (AE) after each piece (Experiment 1 and 2), while peripheral signals (cardiorespiratory measures, skin conductance, and facial muscle activity) were continuously measured (Experiment 2). Factor scores of AE were significantly higher in the AV condition in both experiments. LF/HF ratio, a heart rhythm that represents activation of the sympathetic nervous system, was higher in the AO condition, suggesting increased arousal, likely caused by less predictable sound onsets in the AO condition. We present partial evidence that breathing was faster and facial muscle activity was higher in the AV condition, suggesting that observing a performer's movements likely enhances motor mimicry in these more voluntary peripheral measures. Further, zygomaticus ('smiling') muscle activity was a significant predictor of AE. Thus, we suggest physiological measures are related to AE, but at different levels: the more involuntary measures (i.e., heart rhythms) may reflect more sensory aspects, while the more voluntary measures (i.e., muscular control of breathing and facial responses) may reflect the liking aspect of an AE. In summary, we replicate and extend previous findings that AV information enhances AE in a naturalistic music performance setting. We further show that a combination of self-report and peripheral measures benefit a meaningful assessment of AE in naturalistic music performance settings.

1 Introduction

There is a clear consensus that listening to music induces aesthetic experiences, with humans augmenting such experiences by optimising the 'where' and 'how' we listen to music, such as in concerts (Sloboda et al., 2012; Sloboda & O'Neill, 2001; Wald-Fuhrmann et al., 2021). Although the aesthetic experience (AE) of music is enhanced in a concert by several aspects (see Wald-Fuhrman et al., 2021 for an overview), one explored here is visual information. While previous work showed that visual cues enhance self-reported musical evaluation of music performances (e.g., see Platz & Kopiez, 2012 for a meta-analysis), some gaps in the literature remain. Firstly, most studies comparing audio-visual (AV) and audio-only (AO) musical performances have been conducted in laboratory settings; to test a more genuine AE, it is imperative to use a more naturalistic situation. Secondly, only two studies so far explored physiological responses between AV and AO musical performances (Chapados & Levitin, 2008; Vuoskoski et al., 2016), but their findings are contrary to each other. Thus, the current study aimed to specify the link between modality (AO vs. AV), AE, and peripheral physiological responses in a naturalistic music performance setting, i.e., a piano concert.

While the initial study of AE has had a strong philosophical focus, AE is currently of great interest in cognitive neuroscience and the neuroscientific subdiscipline of neuroaesthetics. Here, perception, emotion, and appreciation are considered to influence AE (for comprehensive reviews, see Anglada-Tort & Skov, 2020; Brattico & Pearce, 2013; Juslin, 2013; Pelowski et al., 2016; Schindler et al., 2017). Specific to the dynamic nature of music, Brattico and colleagues (2013) proposed that the AE of music listening is composed of a chronometry of components: 1) perceptual sensory processes (feature analysis/integration) as well as early emotional reactions (e.g., startle reflex and arousal), 2) cognitive processes (based on long-term knowledge, such as harmonic expectancy), and 3) affective processing (including perceived and felt emotions). A combination of these processes that involve somatomotor processes interacting with the listener themselves (in terms of cultural knowledge, musical expertise, etc.) and external context (e.g., social setting), result in 4) aesthetic responses (emotions, judgements, and liking). Brattico et al. (2013) presented neurophysiological correlates that might reflect these processes. Namely, sensory processes should be reflected in early event-related potentials (ERPs) and in early auditory processing areas (sensory cortices, brainstem). More cognitive ('error' and 'surprise') components should be reflected in the MMN and P300 and non-primary sensory cortices. Finally, (aesthetic) emotion and judgements should be reflected in the late potential component (LPC) and reward and emotion areas in the brain. Research further suggests that (synchronisation of) certain brain oscillations are related to music-evoked pleasure, particularly frontal theta oscillations (Ara & Marco-

Pallarés, 2020; Sammler et al., 2007; Tervaniemi et al., 2021), parieto-occipital alpha (Nemati et al., 2019), theta (Chabin et al., 2020) and theta phase synchronisation (Ara & Marco-Pallarés, 2020, 2021), as well as the inter-brain synchrony (IBS) of frontal and temporal theta in shared musical pleasure (Chabin et al., 2022).

Although some work has explored music-evoked pleasure with EEG in the more naturalistic setting of a concert hall (Chabin et al., 2020, 2022), measuring brain activity in such settings comes with significant challenges. A more accessible approach, however, has been to measure peripheral physiological responses in naturalistic settings such as theatres (Ardizzi et al., 2020), concert halls (Egermann et al., 2013), and cathedrals (Bernardi et al., 2017). Peripheral measures include the somatic (voluntary muscle) and autonomic nervous systems (ANS), of which the latter comprises the sympathetic ('fight-or-flight') and parasympathetic ('rest-and-digest') nervous systems (SNS, PNS). In naturalistic settings, previous work revealed (synchronised) physiological arousal responses in audiences occur in relation to surprising, emotional, and structural moments in music such as transitional passages, boundaries, and phrase repetitions (Czepiel et al., 2021; Egermann et al., 2013; Merrill et al., 2021). Such peripheral measures are likewise mentioned in the AE chronometry approach (Brattico et al., 2013) as reflecting tension and chill responses (Grewe et al., 2009; Salimpoor et al., 2009). However, unlike brain regions (fMRI) and the latency/polarity of (EEG/MEG) components, that can be attributed to psychological processes (Kappenman & Luck, 2011), peripheral responses are mainly characterised according to increased/decreased activity, making it more difficult to separate responses relating to distinct sensory, cognitive, and/or aesthetic processes. Thus, rather than taking a superficial understanding that such measures directly index a pleasurable experience, a more thorough biological understanding is required to appropriately interpret the meaning of such measures (see e.g., Fink et al., 2023 for an example in pupillometry).

The current dependent measures of interest, which have also previously been used in research on musical aesthetics (e.g., Grewe et al., 2009; Salimpoor et al., 2009), range from involuntary ANS responses to voluntary motoric control, namely: skin conductance, heart, respiratory, and muscle activity. Skin conductance (SC, also known as electrodermal activity, EDA) measures activation of sweat glands, which are innervated by the SNS only. The heart consists of cardiac muscle (involuntary control), with SNS (via sympathetic nerves) and PNS (vagus) innervations that increase and decrease heart rate (HR), respectively. Typically, HR fluctuates and is measured by different heart rate variability (HRV) measures. These measures can be in the time-domain, for example, the standard deviation between interbeat intervals, or in the frequency-domain, for example, power of certain frequency bands related to SNS and PNS activation. Power at a high frequency (HF, 0.4-0.15Hz) component is attributed to PNS activity, while power at a low frequency (LF, 0.04 - 0.15 Hz) component seems to reflect both PNS and SNS influences; thus, the LF/HF ratio is used to represent SNS activity (Shaffer & Ginsberg, 2017; Task Force of the European Society of Cardiology, 1996). Respiratory activity encompasses both involuntary control - where the lungs are innervated by both SNS and PNS, which dilate and constrict the bronchioles, respectively - as well as voluntary control (Purves & Williams, 2001). The somatic (muscle) system consists mainly of skeletal (voluntary) muscle, commonly measured are the facial muscles of zygomaticus major ('smiling') and corrugator supercilii ('frowning'). Although under voluntary control, certain facial muscle responses may be partly unconscious (i.e., occur without attention or conscious awareness, Dimberg et al., 2000). Overall, SC, heart, respiration, and facial muscle activity broadly relate to arousal and valence[9]. Higher arousal has been associated with SNS activation, such as increased sweat secretion, increased LF/HF ratio, HR and RR acceleration, and decreased HF power (Luft & Bhattacharya, 2015; Shaffer & Ginsberg, 2017), while zygomatic and corrugator muscle activity seem to reflect positive and negative valence, respectively (Bradley & Lang, 2000; Cacioppo et al., 2000; Dimberg et al., 2000; Lang et al., 1993; Larsen et al., 2003, though see discussion below).

Although broadly reflecting arousal and valence, peripheral measures have been related to sensory, cognitive, and aesthetic experiences with regard to acoustic/musical stimuli in separate studies. Increased SC and HR patterns have been related to early sensory reactions to an acoustic signal - referred to as an orienting response/startle reflex (Barry, 1975; Barry & Sokolov, 1993; Graham & Clifton, 1966; Roy et al., 2009). Physiological changes occur in response to cognitive music processes such as recognising unexpected harmonic chords (Koelsch et al., 2008; Steinbeis et al., 2006) and deviant stimuli (in an MMN-like paradigm, Chuen et al., 2016; though see Lyytinen et al., 1992), which might be enhanced by attention (Frith & Allen, 1983). In more naturalistic music listening, many studies showed that arousing music (faster tempi and unpredictable harmony) increase SC, HR, and RR (Bernardi et al., 2006; Coutinho & Cangelosi, 2011; Czepiel et al., 2021; Dillman Carpentier & Potter, 2007; Egermann et al.,

[9] The two main dimensions of emotion, according to the dimensional model of emotion (Russell, 1980). These terms reflect bipolar continuums: arousal ranging from calm to excitement, while valence varies from negative to positive emotional experience. Such peripheral responses have also been attributed to the discrete (basic) emotion theory, where SNS activation relates to happiness/fear, while PNS activation relates to calmness /sadness). For a more thorough discussion on emotion models, see for example (Barrett & Russell, 2015; Hamann, 2012).

2013; Khalfa et al., 2002; Krumhansl, 1997), though we note this result is not consistent across studies, for reviews see Bartlett, 1996; Hodges, 2009; Koelsch & Jäncke, 2015). In terms of valence, researchers have shown that zygomaticus activity increases during happy music (Lundqvist et al., 2008). However, other work showed it can increase during unpleasant (dissonant) music (Dellacherie et al., 2011; Merrill et al., 2021). This conflict suggests that perhaps the activation of the smiling muscle is not just related to valence (see also Wingenbach et al., 2020). Peripheral responses have likewise been related to aesthetic experience of music, or least music-evoked “chills” (frissons), which increases SC, HR, RR and EMG (Blood & Zatorre, 2001; Craig, 2005; Grewe et al., 2009; Salimpoor et al., 2009). Hence, evidence suggests that peripheral measures can reflect (a mixture of) the sensory, cognitive and/or preference parts of the AE, rather than being a direct index of AE. Therefore, it is of importance to collect self-report measures to further interpret the peripheral responses to AV and AO comparisons.

In terms of modality effects on self-reports, audio information seems to be consistently influenced by performer movement. In one percussion study, pairing visual gestures that created long notes to acoustic sounds of short notes resulted in short sounds being perceived as longer sounding notes (Schutz & Lipscomb, 2007); an effect later shown to be consistent in percussive (but not sustained) sounds when the sound appears after a gesture (Schutz & Kubovy, 2009). In piano performances, one acoustic performance was paired with four videos: one as the original performance and three pianist ‘doubles’. Ninety-two out of ninety-three participants perceived differences between the performances, although the sound remained identical (Behne & Wöllner, 2011). With regard to more aesthetic influences, several studies that compared uni- and bimodal versions of music performances showed visual cues enhance a listener’s perception of performance quality (Waddell & Williamon, 2017), musical expertise (Griffiths & Reay, 2018; Tsay, 2013), musical expression (Broughton & Stevens, 2009; Davidson, 1993; Lange et al., 2022; Luck et al., 2010; Morrison & Selvey, 2014; Vines et al., 2011; Vuoskoski et al., 2014), perception of emotional intention (Dahl & Friberg, 2007; Vines et al., 2006), and felt emotion (Van Zijl & Luck, 2013). As AE is related to the appreciation of performance expressiveness, quality, and emotion (Brattico & Pearce, 2013; Juslin, 2013), this research, as well as a meta-analysis (Platz & Kopiez, 2012), showed that AE increases with additional visual cues. One neuroaesthetic theory that could further explain this enhanced AE postulates that visual information may increase embodied simulation, which subsequently increases AE (Freedberg & Gallese, 2007; Gallese & Freedberg, 2007). Support for this idea comes from studies showing higher activation in the action observation network when viewing movements that are rated as aesthetically pleasing (Cross, 2011).

However, this enhanced AE effect has been mostly assessed in laboratory settings. Recent studies are increasingly exploring such experiences in live concerts (Chabin et al., 2022; Coutinho & Scherer, 2017b; Czepiel et al., 2021; Scherer et al., 2019a; Swarbrick et al., 2019; Tervaniemi et al., 2021), where participants report experiencing stronger emotions (Gabrielsson & Wik, 2003; Lamont, 2011); however, Belfi et al. (2021) found that felt pleasure did not differ between live and an audio-visual recording of the same performance. Focusing more specifically on the role of modality, to date only a few studies compare responses to AV vs. AO conditions in naturalistic settings. Compared to eyes-closed conditions, eyes-open conditions increased movement energy and interpersonal coordination, suggesting that visual information may enhance the social aspect of live pop/soul music (Dotov & Trainor, 2021). Coutinho & Scherer (2017b) compared emotional responses in a live AV performance to recorded AV, AO, and VO performances of Schubert Lieder, where the live AV condition had significantly higher wonder and significantly lower boredom ratings. Although these two studies highlight the difference between genres and the affordances that visual information can give (focus on seeing other audience members/musicians in popular/classical music, respectively), they essentially show that additional information enhances the (social/emotional) experience. We stress that it is not trivial to replicate findings from the lab to a more naturalistic setting, since, for example, well documented effects of familiarity and body movement on music appreciation found from lab studies were not replicated in a field study (Anglada-Tort et al., 2019). It is also worth extending Coutinho & Scherer (2017b), since they focus on the more emotional part of AE, and only collected data from an AV modality in a naturalistic setting (other modalities were tested in a lab-like setting). The current study thus compares modalities in one naturalist setting to examine more specifically the judgement and preference components of AE.

Two previous studies have compared peripheral physiological responses as a function of modality during music performances and serve as the starting point for the current work. Chapados & Levitin, (2008) found that self-reported tension as well as SC were both highest in AV conditions. However, Vuoskoski et al. (2016) found that, although self-reported intensity, high energy arousal, and tension were highest in AV conditions, SC was actually highest in AO conditions. While the discrepancy between these two studies could relate to the different styles and instruments used (which offer different expressive affordances), Vuoskoski et al. (2016) argued that SC might be higher during AO performances due to musical expectancy (Huron, 2006; Juslin & Västfjäll, 2008). More specifically, as visual information increases listeners' ability to predict upcoming musical events, AV stimuli are less surprising. Indeed, this idea is supported by speech studies focusing on the N100, an EEG event-related potential component that reflects early sensory processing, where a larger N100 amplitude can indicate a response to a less predictable stimulus. The N100 component is enhanced in AO (compared to AV) conditions in speech (Klucharev et al., 2003; van Wassenhove et al., 2005), emotional expression (Jessen & Kotz, 2011), as well as non-speech events such as clapping (Stekelenburg & Vroomen, 2007). These findings corroborate the idea that the lack of visual information makes sound onsets less predictable.

Together, this evidence suggests that peripheral responses might be 1) higher in AO conditions if they reflect sensory processing, or 2) higher in AV conditions if they reflect the enhanced emotional and/or appreciation aspects of AE. If peripheral physiological responses reflect sensory processing, we expected to replicate results from Vuoskoski et al. (2016) and find increased physiological activity in AO conditions. However, if physiological responses reflect the more emotional/aesthetic aspects, we expected to replicate results from Chapados & Levitin (2008) and find increased physiological responses in AV conditions.

In summary, more research is needed to assess modality effects that enhance aesthetic experience in a more naturalistic setting. Further, the peripheral physiological correlates of aesthetic effects are so far inconsistent. The current study consists of two experiments that examine AE and physiology between AV and AO conditions in a concert hall setting. In both Experiments, we recorded behavioural responses and tested the hypothesis that AE will be higher in the AV condition. In Experiment 2, we additionally collected physiological responses and tested the hypothesis put forward by Vuoskoski et al. (2016) that peripheral physiological activity should be higher in AO conditions.

2 General Method

Overview

We present two experiments, each consisting of two concerts. Experiment 1 (Concerts 1 and 2) measured behavioural ratings, while Experiment 2 (Concerts 3 and 4) measured both behavioural ratings and physiological responses. Both involve the same stimuli and the same within-subjects experimental design: participants listening to piano performances of Bach, Beethoven, and Messiaen, in AO and AV conditions. Modality order was counterbalanced across concerts.

Stimuli

Upon engaging a pianist, three musical pieces were selected from their repertoire in accordance with the pianist and musical experts to represent various emotional expressions (cheerful, sad, and ambiguous) and musical styles (Baroque, Classical-Romantic, and 20th century music): Johann Sebastian Bach: Prelude and Fugue in D major (Book Two from the Well-Tempered Clavier, BWV 874), Ludwig Van Beethoven: Sonata No. 7, Op. 10, No. 3, second movement (Largo e mesto), and Olivier Messiaen: *Regard de l'Esprit de joie* (No. 10 from Vingt Regards sur L'Enfant-Jésus). These pieces were presented to the participants during each concert twice in the two different modalities: in audio-visual (AV) and an audio-only (AO) versions. We considered this repetition of pieces as a naturalistic part of the design as piece repetition is a practice (although not extremely common) in concert programming (Halpern et al., 2017).

Both AV and AO presentations of the music pieces were performed by the same pianist, playing on the same piano (Steinway B-211), in the same concert hall. AV versions of the music pieces were performed live during the concerts and the audience could see and hear the pianist performing the music. AO versions of the music pieces were recorded in the same concert hall, on the same piano in advance of the concerts, without an audience. The AO versions were presented during the concerts via a stereo setup with high-quality full-range loudspeakers (Fohhn LX-150 + Fohhn XS-22), so that the audience could only hear the music. During this time, the pianist was backstage, so that the audience could only see the piano. The playback AO versions were the same in all concerts in both experiments. To ensure similarity of sound levels between AO and AV presentations, a trained sound engineer checked that the loudness across the modalities was equal.

Although modality conditions were controlled as much as possible, we would assume that repeated performances of the same musical piece might have slight deviations from each other, even when performed by a highly trained professional musician (Chaffin et al., 2007). Therefore, we checked that the stimuli nonetheless were comparable enough to eliminate confounding variables of potential acoustic differences between AV versions (different for each concert) and AO versions (the same across all concerts). We differentiated between score-based features and performance-based features (Goodchild et al., 2019). The former refers to features that come from the notated scores (e.g., harmonies), which should remain the same across performances (assuming no errors in playing the scores). The latter refers to features that may also be notated in the scores (e.g., dynamic markings) but might deviate more depending on the performances, such as tempo, loudness, and timbre. Tempo was extracted using a combination of MIDI information for each note and manually locating the beat (using Sonic Visualiser, Cannam et al., 2010), where inter-beat intervals were obtained to calculate continuous beats per minute (bpm). Loudness and timbre were extracted from the audio signal using MIRToolbox (Lartillot & Toiviainen, 2007) in MATLAB, with RMS (*mirrms*) and spectral centroid (*mircentroid*) representing loudness and timbre, respectively. In checking multicollinearity (Lange & Frieler, 2018), none of the features correlated highly, confirming that each feature represented an independent aspect of the music. Each of the features were averaged into average bins per bar (American: measure) to account for slight timing deviances between performances. The features over time were very similar (see Supplementary Figures 1-6 in Supplementary Materials). This similarity was confirmed by significant correlations between concerts, all with *r* values > 6 (see Supplementary Materials, Supplementary Table 1), suggesting that all performances were acoustically comparable.

Questionnaires

Questionnaires were presented after each musical piece to assess three types of questions. Firstly, we assessed the 'naturalness' of the concert by asking to what extent the experimental components of the setting (e.g., measurement of the behavioural responses) disturbed the concert experience, where 'disturbed by measurement' was rated from 1 (strongly disagree) to 7 (very much agree). We further assessed familiarity with the style of music as well as whether the participant knew the specific piece of music. This was rated from 1 (not at all familiar) to 7 (very familiar). Thirdly, we assessed the main dependent variable of interest: aesthetic experience (AE). As an AE is made up of several components (Brattico & Pearce, 2013; Schindler et al., 2017), we assessed the aesthetic experience with a set of eight individual items, consisting of how much they liked the piece, how much they liked the interpretation of the piece, and how absorbed they felt in the music (see Supplementary Materials for all questions).

Procedure

Participants were invited to attend piano concerts that took place at the ArtLab of the Max Planck Institute for Empirical Aesthetics in Frankfurt, a custom-built concert hall seating 46 audience members (https://www.aesthetics.mpg.de/en/artlab/information.html). Concerts were kept as identical as possible for factors such as lighting, temperature, and timing. Prior to the concert, participants were informed about the experiment and filled in consent forms before being seated in the ArtLab. During the concert and after each piece of music, participants answered the short questionnaire described above. All participants saw the three pieces in both conditions. For one concert per Experiment, the three music pieces were presented first in the AO modality, and then repeated in the AV modality. Modality order was counterbalanced so that in the other concert per Experiment, music pieces were first presented in the AV modality, and then again in the AO modality. An overview of the procedure and modality condition orders can be found in Figure 1. Behavioural measures were recorded in both

Experiment 1 and 2. In Experiment 2 only, physiological data were additionally collected (details in Section 4.1.2 Experiment 2 Procedure).

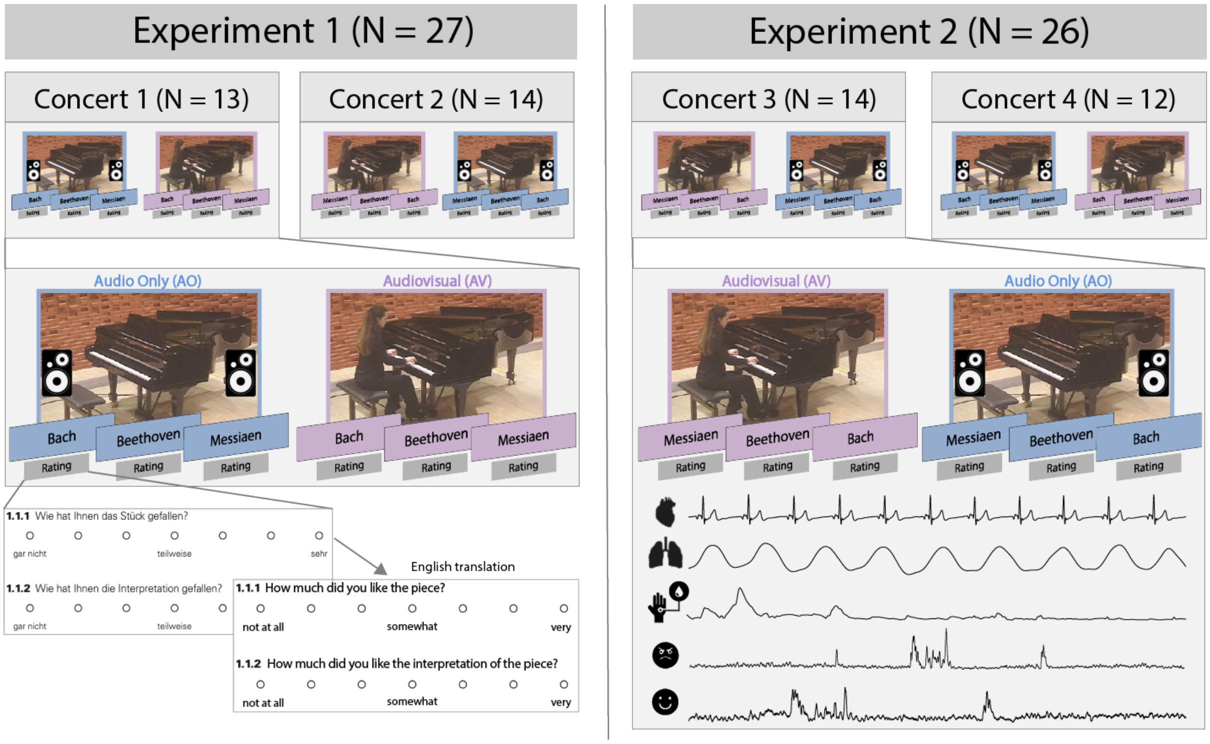


Figure 1. Outline of the experimental procedure in Experiment 1 (behavioural audience ratings) and Experiment 2 (audience ratings and peripheral physiological measures). Pieces were presented both in an AV version (purple boxes) and an AO version (presented via speakers, blue boxes).

Analysis

Statistical analyses were conducted in R and R studio (R Core Team, 2021; RStudio Team, 2021).

Items chosen for the questionnaire (see Supplementary Materials) reflect elements of an aesthetic experience. Thus, it was assumed that the items might be related to each other. Indeed, in both Experiments, items in the self-reports capturing the aesthetic experience (scaled within participants, using *scale* function in R) were highly correlated. Therefore, rather than comparing modality differences for each item, we reduced the questionnaire items to an overall, more interpretable factor - that retains important information from each item - using a factor analysis (Fabrigar et al., 1999). This reduced factor yielded new factor scores that mixed scores from the original items together based on loadings, i.e., regression weights (using *fa* from the psych package, see accompanying code; Revelle, 2022). The more one item contributed to - or loaded onto - the reduced factor, the higher the 'item loading' was for that factor. Table 1 shows the item loadings of factors in both experiments. These factor scores were used as a new overall variable that represents a summary of the questionnaire items. Details about each FA for each experiment are explained below in the experiment-specific methods.

Linear mixed models (LMMs) were run with the factor scores extracted from the factor analysis as the dependent variable, with modality (AV / AO) as a predictor (fixed effect). We also ran LMMs for each physiological measure, where modality was the predictor (fixed effect) as well as a LMM assessing relationship between factor scores and physiological measures. LMMs are more appropriate than repeated measures ANOVA, as they are more fitting for physiological data, can account for missing trials, and can model random sources of variance and non-independence in the observations (Barr et al., 2013; Page-Gould, 2016; Winter, 2013). Ratings and physiological measures were recorded multiple times from each participant, who heard the same music piece more than once, in groups for each concert. To account for these random sources of non-independence, we added random intercepts for concert, piece, and participant. Participants were nested within concerts, while participant and piece were considered crossed effects. For the physiological data, piece sections were further nested within pieces to account for observations taken within pieces (see Methods for Experiment 2). We also included a random slope for participants. Thus, the models represent the maximal random effects structure justified by the design (Arnqvist, 2020; Barr, 2021; Barr et al., 2013). While LMMs do not rely on normally distributed data, we checked linearity, homoscedasticity, and normality of residuals of the models (Winter, 2013). We also checked for model errors. All maximal models generated singular fit errors, suggesting that the model might be too complicated and/or one or more random effects have (near to) zero variance or (near-)perfect correlations. Therefore, we followed the recommended procedure of simplifying models until error is removed (Barr, 2021), ultimately selecting a model with a random effect structure that is supported by the data (Barr et al., 2013; Matuschek et al., 2017). As error-free models are generally preferred (Barr et al., 2013), we report the models that generated no errors, but report all maximal and simplified models in the Supplementary Materials. LMMs were run using *lmer* from the *lme4* packages (Bates et al., 2015; Kuznetsova et al., 2017). Significance values, effect sizes, and Akaike information criterion (AIC) were obtained from the *tab_model* function from *sjPlot* package (Lüdecke, 2023). Pairwise comparisons were run with the *emmeans* function from *emmeans* package (Lenth, 2021) with Bonferroni corrections. As a sanity check for the linear mixed models, we also ran ANOVAs (Arnqvist, 2019). Corresponding code and required to run these analyses are available at Open Science Framework (OSF): (Please note this repository is currently private and only available with this link while the manuscript is under review; it will be made public when the manuscript is accepted).

3 Experiment 1

Method

Participants

The study was approved by the Ethics Council of the Max Planck Society and in accordance with the declarations of Helsinki. Participants gave their written informed consent. Twenty-seven participants attended the experimental concerts (13 and 14 participants in Concert 1 and 2, respectively), 18 females (9 males), with mean age of 57.96 years (SD = 20.09), who on average had 6.99 years of music lessons (SD = 7.87) and attended approximately 13 concerts in the last 12 months (M = 12.62; SD = 13.37). Participants also provided ratings on their perception being a musician (from 1 = does not apply, to 7 completely applies), most participants selected 1 (N = 13) or 2 (N = 4), and less selected 3 (N = 1), 4 (N= 2), 5 (N = 3), 6 (N = 2) and 7 (N = 2). Most had a college/university degree (N = 22), the others either vocational training (N = 2) or completed A-levels/high school (N = 3). Wilcoxon tests showed that participants did not differ in Concert 1 and 2 in terms of age ($p = .590$), musician level ($p = .877$), years of music lessons ($p = 1.00$), and number of concerts attended in the last 12 months ($p = .173$).

Factor analysis and statistical analysis

Questionnaire items were chosen to reflect elements of an aesthetic experience. As they were highly correlated (see accompanying code), we chose to reduce these variables to an interpretable factor using factor analysis. A Kaiser-Meyer-Olkin (KMO) measure verified sampling adequacy (KMO = .801, well over the .5 minimum required) and all KMO values for individual items were > .670. Bartlett's test of sphericity was significant, revealing that correlations between items were large enough for a FA, $X_2(28) = 408.844$, $p < .001$. Kaiser's criterion of eigenvalues > 1 and a scree plot indicated a solution with one factor. Thus, a maximum-likelihood factor analysis was conducted with one factor, which explained 37% of the variance. We took the scores of this factor and created a new variable. As items of liking, liking of interpretation, and absorption loaded highly onto this factor, and these aspects have been identified as critical aspects of an aesthetic experience (Brattico & Pearce, 2013; Orlandi et al., 2020), we referred to this new variable as the overall 'aesthetic experience' (AE). Nine trials with an outlier exceeding ±3 Median Absolute Deviations (MAD, Leys et al., 2013) was removed from further analyses. In total, we had 153 observations for the AE scores [(27 participants x 3 pieces x 2 modality conditions) - 9]. We compared AE factor scores between modality conditions using LMMs (see General Methods, corresponding code).

Table 1. FA loadings from questionnaire items in both Experiment 1 and 2. Factor 1 for both Experiment 1 and 2 is interpreted as 'Aesthetic experience'.

	Experiment 1	Experiment 2
Item	Factor 1	Factor 1
Liking	0.78	.87
Liking of interpretation	0.69	0.63
Absorption	0.90	0.67
Passive reception	0.73	0.09
Connection to musicians	0.56	0.43
Urge to move	0.17	0.21
Connection to co-listeners	0.28	0.14
Understanding	0.27	0.34

Results

Assessing naturalistic situations.

Results of whether the measurements disturbed the concert or not are shown in Table 2. The mean rating was 1.537 (SD = 1.016) out of 7, with 88% of ratings at 1 or 2 on the scale (i.e., strongly disagree or disagree that measurements disrupted the concerts, respectively). Thus, behavioural measurements did not disrupt the concert, confirming the ecological validity of the experimental setting.

Piece familiarity.

Ratings for familiarity of style were similarly high for Bach (M = 5.796, SD = 1.279) and Beethoven (M = 5.630, SD = 1.248), but lower for Messiaen (M = 3.333, SD = 1.981). Most participants did not know the pieces specifically, though 18%, 26%, and 11% of participants knew the Bach, Beethoven, and Messiaen pieces, respectively.

Aesthetic experience: Modality differences.

LMMs showed modality was a significant predictor of AE (see Table 4). AV scores were significantly higher (M = 0.186, SE = .296, 95% CI [-0.962 1.33]) than AO scores (M = -0.102, SE = .297, 95% CI [-1.245, 1.04]), $t(124)= -.240$, $p = .018$) (see Figure 2). This effect was confirmed by the maximal model, despite generating a singular fit error: it yielded the same estimates and had similar effect sizes, AIC, and significance (see Supplementary Table 3). The modality effect was confirmed by an ANOVA, which yielded a significant main effect of modality ($F(1,26) = 5.564$, $p = .026$).

Table 2. Ratings of feeling disturbed by the measurement, and familiarity with style and specific piece in Experiment 1.

Ratings of feeling disturbed by the measurement							
Rating	1	2	3	4	5	6	7
	69%	19%	5%	3%	3%	1%	0%
Familiarity with style of piece							
Rating	1	2	3	4	5	6	7
Bach	0%	2%	4%	11%	18%	26%	39%
Beethoven	0%	2%	4%	15%	16%	35%	28%
Messiaen	26%	17%	9%	20%	9%	11%	8%
Familiarity with specific piece							
	0 (No)		1 (Yes)			Not sure	
Bach	78%		18%			4%	
Beethoven	70%		26%			4%	
Messiaen	85%		11%			4%	

Discussion

Experiment 1 tested whether participants had higher AE in the audio-only (AO) or audio-visual (AV) piano performances in a naturalistic concert setting. We confirmed that the measurements did not disturb participants and the findings show that AE increased in the AV compared to AO condition. These results support prior experimental laboratory results that showed liking and appreciation of expressivity are increased in AV conditions (Platz & Kopiez, 2012). We confirm that these results can be extended in a more naturalistic setting. One study that compared emotional differences between modalities in a naturalistic context, found higher wonder ratings but lower boredom ratings in live AV performances of music (Coutinho & Scherer, 2017b). Our results likewise fit and extend this work, showing that the preference (liking) and absorption of the AE is also higher in AV modality. As naturalistic environments allow less control, it is important that these findings are replicated.

4 Experiment 2

Previous studies aimed at gaining further insight into potential emotional differences between uni- and bimodal music performances by measuring physiological responses (Chapados & Levitin, 2008; Vuoskoski et al., 2016). However, so far results are inconsistent. In Experiment 2, we explored whether different modalities would affect peripheral physiological responses similarly to the behavioural responses of AE (Exp. 1), and whether one or the other peripheral signal might better serve as an index of AE.

Method

Participants

The study was approved by the Ethics Council of the Max Planck Society and in accordance with the declarations of Helsinki. Participants gave their written informed consent. Twenty-six participants in total attended either Concert 3 (N=14) or Concert 4 (N = 12). Experiment 2 in total included nine females (17 males), with a mean age of 51.64 years (SD = 15.41), who on average had 5.94 (SD = 8.13) years of music lessons and attended an average of 14 concerts per year (M = 13.62, SD = 19.70). Participant provided ratings on their perception being a musician (from 1 = does not apply, to 7 completely applies), and most participants selected 1 (N = 15) or 2 (N = 3), while less selected 3 (N = 0), 4 (N= 1), 5 (N = 4), 6 (N = 2), or 7 (N = 1). All had either vocational training (*N* = 7) or a college/university degree (*N* = 19). Wilcoxon tests showed no significant differences between participants in Concert 3 and Concert 4 in terms of age (p = .72), years of music lessons (p = .14), and number of concerts attended in the last 12 months (p = 1.00). There was a significance in musician level between concerts (p = .039).

In assessing differences between the participant samples of the two Experiments, Experiment 1 had a significantly older audience on average (mean age in Experiment 1 = 58, Experiment 2 = 52, p = .041), but no significant differences for number of music lessons (p = .334), concert attendance in the last 12 months (p = .755), and musician level (p = .575).

Self-report data from all 26 participants were used in the analysis, while one physiological dataset from Concert 3 was lost due to technical problems (physiology: N = 25).

Procedure

Participants were invited to arrive an hour before the concert, during which they were fitted with physiological equipment. All signals were collected with a portable recording device, 'plux' (https://plux.info/12-biosignalsplux), that continuously measured physiology across the duration of the concert at a 1000 Hz sampling rate. Respiration was measured via two respiration belts: one respiration belt was placed around the upper chest of the participant, and one respiration belt was placed around the lower belly. ECG, EMG, and EEG were collected using gelled self-adhesive disposable Ag/AgCl electrodes. Locations for the EMG, EEG, and ECG were prepared with peeling gel (under the left eyebrow and on left cheek for EEG, on the chest for ECG, and on the forehead for EEG). Three ECG electrodes were placed on the chest in a triangular arrangement; two as channels and one as the ground. Two facial muscles were recorded on the left side of participants' faces; two electrodes were placed at the zygomaticus major ('smiling') muscle, and two electrodes were placed on the corrugator supercilii ('frowning') muscle, with a ground placed behind the left ear. EDA was collected via two electrodes placed on the middle phalanges of the non-dominant hand of participants. EEG activity from the frontal region was collected from three electrodes placed on the upper forehead, with a reference electrode placed in the middle of the forehead (in a similar location to an Fpz location in a conventional EEG cap), with additional two electrodes placed above the left and right eyebrows (in a similar position to Fp1 and Fp2 in a conventional EEG cap, respectively). EEG data are not reported in this paper.

Factor analysis

We used the same items as in Experiment 1. Again, these item ratings were highly correlated (see accompanying code) and we chose to reduce these variables with a factor analysis. A Kaiser-Meyer-Olkin measure verified sampling adequacy (KMO = .609). All but one item had KMO values > .5; this one item ('connection with co-listeners') had a value of close to .5 (0.416). Correlations between items were large enough for a FA (Bartlett's test of sphericity, $X_2(28) = 264.725$, $p < .001$. Kaiser's criterion of eigenvalues > 1 and a scree plot indicated a solution with one factor. Thus, a maximum likelihood factor analysis was conducted with one factor, which explained 24% of the variance. We took the scores from this factor and created a new variable. As we had similar loadings to Experiment 1, we also refer to this factor as the overall aesthetic experience (AE). In this factor, eleven outlier values exceeding ±3 Median Absolute Deviations (MAD, Leys et al., 2013) were removed from further analyses. In total, we had a total of 145 observations [(26 participants x 3 pieces x 2 modality conditions) - 11].

Physiological pre-processing.

Pre-processing of physiological signals (Experiment 2) was conducted in MATLAB (2019b, The Mathworks Inc, USA). Any missing data (gaps ranging from 5 - 53 ms long) were first linearly interpolated at the original sampling rate. Continuous data were then cut per piece. Using Ledalab (www.ledalab.de), skin conductance data were manually screened for artefacts (8% of data were rejected), downsampled to 20 Hz and separated into phasic (SCR) and tonic (SCL) components using Continuous Decomposition Analysis (Benedek & Kaernbach, 2010). Following previous literature, data were detrended to remove remaining long-term drifts (Omigie et al., 2021; cf. Salimpoor et al., 2009). Respiration, ECG, and EMG data were pre-processed using the Fieldtrip (Oostenveld et al., 2011) and biosig toolboxes in MATLAB (http://biosig.sourceforge.net/help/index.html). Manual screening of respiration data showed that the respiration signals obtained from the lower belly were stronger than those obtained from the upper chest; only data from the respiration belt around the lower belly were therefore used for further analysis. Respiration data were low pass filtered at 2 Hz, ECG data were band-pass filtered between 0.6 and 20 Hz (Butterworth, 4th order), and both demeaned. QRS peaks in the ECG signal were extracted using *nqrsdetect* function from biosignal, and peaks were found in respiration using custom functions. Computationally identified peaks were manually screened to ensure correct identification; any missing QRS peaks were manually added, while falsely identified QRS peaks were removed. Any ECG/respiration data that were too noisy for extraction of clear QRS/respiration peaks were rejected from further analysis (ECG = 14%, respiration = 7%). Differential timing of signal peaks – i.e., interbeat intervals (IBI, also known as RR-intervals) for ECG, and inter-breath intervals (IBrI) for respiration – were converted to beats per minute and interpolated at the original sampling rate to obtain a continuous respiration and heart rate. Heart rate variability measures were extracted using the *heartratevariability* function in biosig (http://biosig.sourceforge.net). Normalised units of high frequency (HF, 0.15–0.4 Hz) power as well as the LF/HF ratio were taken into further analysis to reflect SNS and PNS activity (frequencies that adhere to the recommendations (Task Force of the European Society of Cardiology, 1996). Electromyography (EMG) data for zygomaticus major (EMGZM) and corrugator supercilii (EMGCS) were band-pass filtered between 90 and 130 Hz and demeaned. We proceeded with the smoothed absolute value of the Hilbert transformed EMG signals.

Although there are questions as to what the most appropriate (central tendency) representation of physiological data is, we relied most closely on the methodology applied by Vuoskoski et al. (2016) to compare results. Therefore, the average of each (pre-processed) physiological measure was the main metric. As physiological responses change over time (i.e., they are non-stationary), and to gain a better representation (signal-to-noise ratio) of the responses across the course of each long piece, data for each piece were divided into piece sections that were driven by the musical structure (which were confirmed by a music theorist). Responses were averaged across these sections. Beethoven was split into nine, Messiaen into nine, and Bach into seven sections (see Supplementary Materials for more information). Overall, we were interested in eight physiological measures: averages of SCL, SCR, HR, HF power and LF/HF ratio, RR, as well as zygomaticus and corrugator activity, which we averaged per participant, modality, piece, and section. As with behavioural data, we removed outliers exceeding ±3 MAD. Total observations for each physiological measure after exclusion of noisy data and outliers were as follows: EMGCS = 1037, EMGZM = 1082, HR = 1073, HF = 1050, LF/HF ratio = 1041, RR = 1152, SCL = 1066, SCR = 910.

Analysis

Statistical analysis for the AE scores obtained in Experiment 2 were conducted as described in Experiment 1. For the physiological responses, we ran LMMs for each of the physiological dependent variables, with modality (AV/AO) as a fixed factor. As data were clustered within groups for participant, piece section, and concert, we added these as random effects in the model (Barr et al., 2013). To determine if behavioural results were related to peripheral responses, we ran a LMM with aesthetic experience as the dependent variable and the eight peripheral measures (all of which were averaged across piece sections to represent rating per piece and scaled to be included in the same model) and condition as predictors. Random effect represented design-driven maximal were included: random intercepts were added for concert, piece, modality condition, and participant. Participants were nested within concerts, while participant, condition, and piece were considered as crossed effects. Variance Inflation Factors (VIF) were checked using the *car* package (Fox & Weisberg, 2019), confirming that VIFs were below 3.

Results

Assessing naturalistic situations.

We first assessed the extent to which the behavioural/physiological measurements disturbed the overall experience during the concert (i.e., for all pieces/conditions). Ratings suggested that measurements did not disrupt the concert experience, with a mean rating of 2.019 (SD = 1.416) and with 75% of ratings at 1 or 2 on the scale. Results are shown in Table 3. These results provide an important validation that physiological measurements can be used in the concert hall settings without impacting ecological validity.

Piece familiarity.

Similar to Experiment 1, ratings for familiarity of style were high for Bach (M = 5.385, SD = 1.484) and Beethoven (M = 5.333, SD = 1.532), but lower for Messiaen (M = 4.135, SD = 1.879). Approximately a third knew the Beethoven and Bach pieces, whereas only 19% knew the Messiaen piece.

Aesthetic experience: Modality differences.

For the behavioural AE results, LMMs showed modality was a significant predictor of AE (see Table 4) with AV scores significantly higher (M = .222, SE = 0.229, 95% CI [-2.07 2.52]) than AO scores (M = .003, SE = 0.229, 95% CI [-2.28 2.29], $t(119) = -0.207$, $p = .041$) (see Figure 2). Although the maximal model generated a singular fit error, it yielded the same estimate and significance, as well as a similar effect size and AIC to the simplified model that generated no error (see Supplementary Table 4). The modality effect was also confirmed by an ANOVA ($F(1,25) = 6.832$, $p = .015$). These results replicated the behavioural findings of Experiment 1.

Physiological differences between modality

LMM results are presented in Table 5 (see also Figure 3). Modality condition was a significant predictor for LF/HF ratio, which represents SNS activation (higher arousal). Comparison of estimated marginal means indicated that this measure was higher in the AO than the AV condition (Table 6). This effect was consistent in the maximal models (see Supplementary Table 12) and confirmed by an ANOVA ($F(1,21) = 5.393$, $p = .030$).

Modality was a significant predictor for respiration rate (RR) and corrugator muscle activity (EMGCS), with a significant increase in the AV compared to AO condition (see Tables 5 and 6). However, in the maximal models that generated errors, the modality effect was not significant for EMGCS nor RR (see Supplementary Table 5 and 10). Corresponding ANOVAs yielded insignificant results for RR ($F(1,22) = 1.95$, $p = .177$), though EMGCS was almost significant ($F(1,21) = 3.679$, $p = .069$). Due to the inconsistency of results between maximal models that generate errors and models with a simplified random structure that is free of errors, findings of EMGCS and RR are only cautiously interpreted.

Peripheral measures that predict behaviour

In a model where AE was the dependent variable and all peripheral measures were predictors, zygomaticus activity (EMGZM) was a significant predictors of self-reported AE (see Table 7): increased smiling muscle activity was positively associated with AE.

Table 3. Ratings of feeling disturbed by the measurement and familiarity with style and specific piece in Experiment 2

Ratings of feeling disturbed by the measurement								na
Rating	1	2	3	4	5	6	7	
	51%	24%	9%	9%	4%	1%	2%	
Familiarity with style of piece								
Rating	1	2	3	4	5	6	7	
Bach	0%	8%	4%	11%	23%	27%	27%	
Beethoven	0%	8%	4%	17%	15%	27%	27%	2%
Messiaen	8%	15%	19%	14%	15%	15%	14%	
Familiarity with piece								
	0		1			Not sure		
Bach	65%		35%			0%		
Beethoven	67%		33%			0%		
Messiaen	79%		19%			2%		

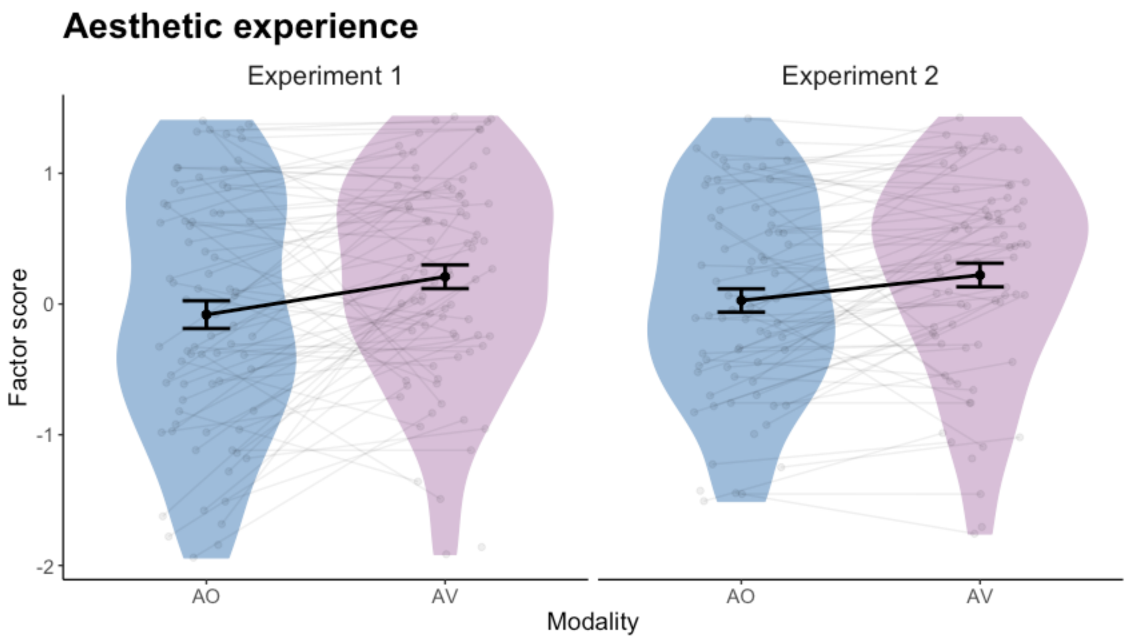


Figure 2. Aesthetic experience factor scores (which had high item loadings of liking, liking interpretation and absorption, see Table 1) as a function of modality (Audio-Only (AO) is blue and Audio-visual (AV) is purple). The left panel shows results for Experiment 1, while the right panel shows results for Experiment 2. Each point represents factor scores for each participant and each piece.

Table 4. Results of Linear Mixed Models comparing Aesthetic Experience factor scores between modality conditions AO and AV.

Aesthetic Experience (AE)

	Experiment 1			Experiment 2		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
(Intercept)	-0.10	-0.69 – 0.48	0.733	0.00	-0.45 – 0.46	0.987
cond [AV]	0.29	0.05 – 0.52	**0.018**	0.22	0.01 – 0.43	**0.040**
Random Effects						
σ^2	0.55			0.40		
τ_{00}	0.04 $_{id\_n:concert}$			0.16 $_{id\_n:concert}$		
	0.24 $_{piece}$			0.08 $_{concert}$		
ICC	0.34			0.37		
N	3 $_{piece}$			2 $_{concert}$		
	15 $_{id\_n}$			16 $_{id\_n}$		
	2 $_{concert}$					
Observations	153			145		
Marginal R^2 / Conditional R^2	0.024 / 0.355			0.019 / 0.383		
AIC	371.967			324.565		

Table 5. Linear mixed models for physiological responses

Physiological results

	EMGCS			EMGZM			HF		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
(Intercept)	0.0025	0.0021 – 0.0029	**<0.001**	0.0022	0.0019 – 0.0024	**<0.001**	0.1384	0.1147 – 0.1621	**<0.001**
cond [AV]	0.0002	0.0001 – 0.0003	**<0.001**	0.0001	-0.0000 – 0.0002	0.180	0.0044	-0.0069 – 0.0157	0.447
Random Effects									
σ^2	0.00			0.00			0.00		
τ_{00}	0.00 section:piece			0.00 section:piece			0.00 section:piece		
	0.00 id_n:concert			0.00 id_n:concert			0.00 id_n:concert		
τ_{11}				0.00 id_n.condAV			0.00 id_n1.condAO		
				0.00 id_n1.condAO			0.00 id_n2.condAV		
				0.00 id_n2.condAV					
ρ_{01}									
ρ_{01}									
ICC	0.65			0.50			0.47		
N	9 section			9 section			9 section		
	3 piece			3 piece			3 piece		
	15 id_n			14 id_n			15 id_n		
	2 concert			2 concert			2 concert		
Observations	1037			1082			1050		
Marginal R^2 / Conditional R^2	0.007 / 0.657			0.003 / 0.502			0.001 / 0.467		

Physiological results (continued 1)

	LF/HF ratio			HR			RR		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
(Intercept)	2.19	1.78 – 2.61	**<0.001**	62.00	54.96 – 69.04	**<0.001**	18.31	17.17 – 19.45	**<0.001**
cond [AV]	-0.26	-0.41 – -0.11	**0.001**	-0.19	-0.44 – 0.05	0.123	0.26	0.09 – 0.43	**0.003**
Random Effects									
σ^2	1.52			4.16			2.11		
τ_{00}	0.00 section:piece			0.11 section:piece			0.30 section:piece		
	0.95 id_n:concert			65.70 id_n:concert			6.71 id_n:concert		
				19.99 concert			0.08 concert		
ICC	0.38			0.95			0.77		
N	9 section			2 concert			2 concert		
	3 piece			9 section			9 section		
	15 id_n			3 piece			3 piece		
	2 concert			15 id_n			15 id_n		
Observations	1041			1073			1152		
Marginal R^2 / Conditional R^2	0.007 / 0.389			0.000 / 0.954			0.002 / 0.771		

Physiological results (continued 2)

	SCR			SCL		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
(Intercept)	-0.00	-0.00 – 0.00	0.156	0.00	-0.02 – 0.03	0.970
cond [AV]	-0.00	-0.00 – 0.00	0.754	0.00	-0.01 – 0.02	0.764
Random Effects						
σ^2	0.00			0.01		
τ_{00}	0.00 section:piece			0.00 section:piece		
	0.00 id_n:concert					
ICC	0.25			0.21		
N	9 section			9 section		
	3 piece			3 piece		
	14 id_n					
	2 concert					
Observations	855			910		
Marginal R^2 / Conditional R^2	0.000 / 0.252			0.000 / 0.213		

Table 6. Results of Linear Mixed Models comparing Physiological responses between AO and AV

Phys	Estimated Marginal Means		Pairwise difference (AO-AV)			
	AO: M, (SE), [95% CI])	AV: M, (SE), [95% CI])	ß	*SE*	*T*	*p*
EMGCS	**0.0025 (0.0002), [0.0022, 0.0029]**	**0.0026 (0.00012) [0.0023, 0.0031]**	**-0.0002**	**0.000**	**-4.423**	**<.0001**
EMGZM	0.0021 (0.0001), [0.0019, 0.0025]	0.0023, (0.0001), [0.0019, 0.0026]	-0.000	0.000	-1.334	.207
HR	62.0 (0.359), [16.5, 107]	61.8 (0.359), [16.3, 107]	0.193	0.125	1.542	.1234
HF	0.138 (0.0012), [0.112, 0.165]	0.143, (0.014), [0.113, 0.172]	-0.007	0.004	-1.927	.054
LF/HF ratio	**2.219 (0.211), [1.76, 2.63]**	**1.93 (0.211), [1.50, 2.37]**	**0.26**	**0.077**	**3.380**	**< .001**
RR	**18.3 (0.587), [11.9, 24.8]**	**18.6, (0.587), [12.1, 25.0]**	**-0.257**	**0.087**	**-2.972**	**.003**
SCL	0.0005 (0.013), [-0.026, 0.027]	0.0027 (0.0013), [-0.023, 0.029]	-0.002	0.007	-0.300	.764
SCR	-0.002 (0.001), [-0.005, 0.001]	-0.002, (0.001), [-0.005, -0.001]	0.0002	0.001	0.313	.755

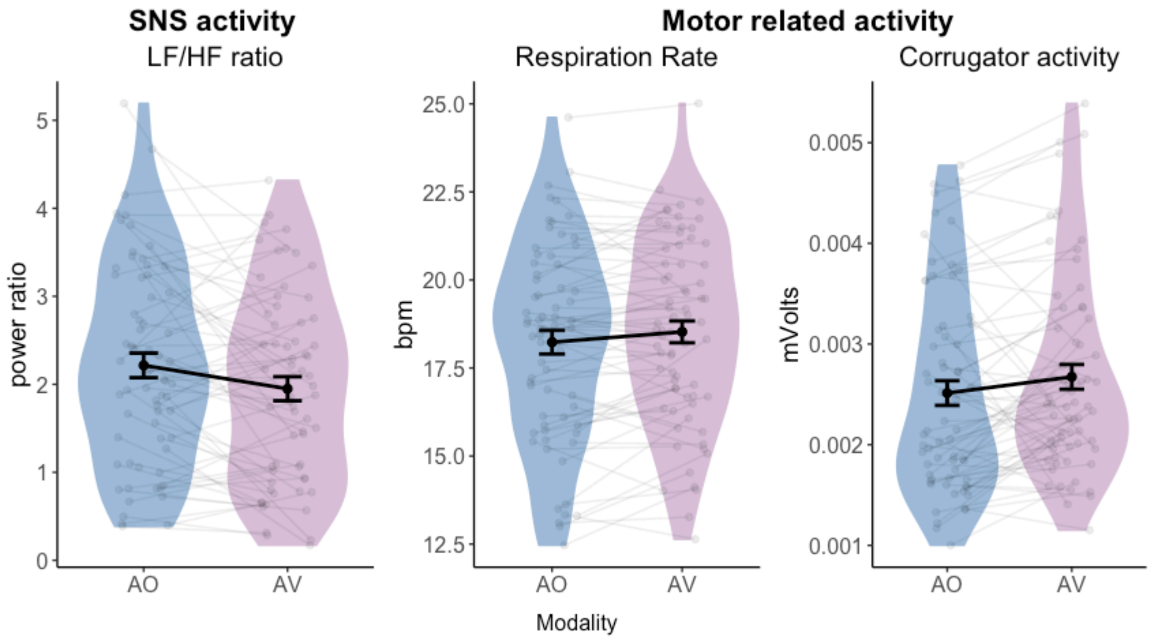


Figure 3. Physiological responses in each modality condition (AO: blue; AV: purple). Different panels represent different physiological measures; from left to right: LF/HF ratio, respiration rate (RR), and EMG activity of zygomaticus major (smiling) muscle (Zygomaticus activity). Each point represents the physiological response value for each participant and each piece.

Table 7. Model of physiology predicting AE self-reports

LMM comparison physiology predicting Aesthetic Experience (AE)

	Maximal model		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	0.1940	-0.1535 – 0.5415	0.270
SCL	0.0982	-0.0737 – 0.2700	0.259
SCR	-0.0108	-0.1884 – 0.1668	0.904
HR	-0.0500	-0.2333 – 0.1333	0.589
RR	0.0628	-0.0984 – 0.2241	0.440
HFnu	-0.0007	-0.2729 – 0.2715	0.996
LFHFratio	-0.0351	-0.3110 – 0.2409	0.801
EMGCS	-0.0767	-0.2508 – 0.0973	0.383
EMGZM	0.1828	0.0196 – 0.3460	**0.029**
Random Effects			
σ^2	0.49		
$\tau_{00\ id\_n:concert}$	0.06		
$\tau_{00\ piece}$	0.00		
$\tau_{00\ cond}$	0.01		
$\tau_{00\ concert}$	0.03		
ICC	0.18		
$N_{concert}$	2		
N_{piece}	3		
N_{cond}	2		
N_{id_n}	13		
Observations	94		
Marginal R^2 / Conditional R^2	0.065 / 0.234		
AIC	256.674		

Discussion

The main aims of Experiment 2 were to replicate the behavioural results of Experiment 1 and to gain further insight into peripheral physiological measures as a function of modality. Importantly, subjective ratings again showed that the measurement of physiological signals did not disturb participants.

As in Experiment 1, AE was significantly higher in the AV condition. We further tested whether peripheral responses between modality conditions. The current study showed that, compared to the AV condition, the AO condition evoked higher LF/HF ratio responses. These results support findings of (Vuoskoski et al., 2016), who reported higher physiological arousal in AO musical performances. On the other hand, respiration rate and corrugator muscle activity, were higher in the AV condition. As both respiration and facial muscle activity are under voluntary muscle control, one interpretation is that viewing movements of the musician increased motor simulation. This is supported by research showing that viewing effortful movements increases respiration (Brown et al., 2013; Mulder et al., 2005; Paccalin

& Jeannerod, 2000) and corrugator activity (de Morree & Marcora, 2010). However, inconsistencies occurred for RR and EMGCS in maximal LMMs compared to error-free LMMs. This model discrepancy suggests the modality effect in respiration and facial muscle activity needs to be complemented and confirmed by further studies with larger sample sizes.

When relating physiological responses to self-reported AE, zygomaticus activity was the only significant predictor. However, as zygomaticus activity increases have likewise been related to unpleasant experiences of (dissonant) music (Dellacherie et al., 2011; Merrill et al., 2021), we only cautiously attribute such facial muscle activity with positive AE.

5 General Discussion

The current experiments aimed to broaden our understanding of naturalistic concert experiences by testing whether (1) AV information enhances aesthetic experience (AE) in a more ecological setting and (2) peripheral physiological responses are higher in AO or AV modality. We also (3) assess the relationship between AE and peripheral physiological responses. We confirm that in both experiments, the measurement of self-report and physiology did not disturb the audiences, supporting the idea that a semi-experimental setting with naturalistic stimulus presentation can yield results of high ecological validity.

As there are several aspects that can make up an AE (Brattico & Pearce, 2013; Juslin, 2013; Schindler et al., 2017), questionnaire items related to certain aspects of an aesthetic experience were used. In both experiments, these items could be reduced to one factor in a factor analysis. Although the factor had slightly different loadings in the two experiments, three main items consistently loaded highly: absorption, liking, and liking of interpretation. Indeed, liking is a strong element of aesthetic experience both in philosophy (as the *evaluative dimension* of AE, Shusterman, 1997) and in empirical work (Brattico & Pearce, 2013). Preference of interpretation (e.g., how fast or expressive) has likewise been shown to play a strong role in AE. For example, observers prefer an expressive – compared to a non-expressive - interpretation of dance (Christensen et al., 2021). Similarly, dance choreography performed with more varied velocities was rated as more aesthetically pleasing compared to when it is performed with a more uniform velocity (Orlandi et al., 2020). Absorption has also shown to be an important factor in mediating aesthetic experience (Brattico & Pearce, 2013) and can even be indexed by peripheral measures, such as microsaccades (Lange et al., 2017). As these items have a strong connection to AE, it seemed appropriate to refer to this factor as such. Further, the fact that all of these items were correlated with each other and captured well by one factor, corroborates previous research that an aesthetic experience comprises many aspects (Brattico & Pearce, 2013; Merrill et al., 2021) and supports the use of dimensionality reduction techniques which trade specificity in favour of a more holistic AE measure.

Both Experiment 1 and 2 consistently showed that AE increases more in the AV than AO modality consistently across models and ANOVAs. Previous laboratory work has revealed that visual information carries several cues of musical expression (Davidson, 1993; Luck et al., 2010), quality (Tsay, 2013; Waddell & Williamon, 2017) and emotion (Dahl & Friberg, 2007; Van Zijl & Luck, 2013), which enhances aesthetic appreciation (Platz & Kopiez, 2012). Though these findings show that AE was significantly higher in AV than AO music performances, the effect size (just under 0.1) was relatively small (Cohen, 1988), likely due to the small sample size. Nonetheless, the overall model effect size (0.3-0.4) is considered medium (Cohen, 1988).

The current results extend the effect of modality influencing musical appreciation in a naturalistic performance setting. Similar work in a concert setting found that the live AV condition had increased wonder and decreased boredom (Coutinho & Scherer, 2017b). However, their main focus was on emotion; we extend their findings to the preference (liking, liking of the interpretation) aspect of AE. We emphasise the importance of conducting AE research in a naturalistic performance setting, as it is more likely to elicit stronger and more realistic responses (Gabrielsson & Wik, 2003; Lamont, 2011). Of note is that results found in laboratory settings are not always replicated in more naturalistic settings. For example, previous laboratory studies have demonstrated that body movement (Platz & Kopiez, 2012) and familiarity (see North & Hargreaves, 2010) increase appreciation of music, even though the latter component has an inverted U-relationship. However, these findings were not replicated in a field study that was conducted in a more realistic situation (busking) and using a dependent variable of appreciation (i.e., money rather than ratings, Anglada-Tort et al., 2019), suggesting that components of music performance influence music appreciation differently depending on the context. Overall, despite the fact that a naturalistic setting might allow less control, together with results from previous work (Coutinho & Scherer, 2017b), we provide consistent support that audio-visual information enhances AE; a finding that likely generalises to more naturalistic human behaviour.

We further elucidated peripheral responses of AE in multimodal contexts (Experiment 2), as research to date is inconsistent (Chapados & Levitin, 2008; Vuoskoski et al., 2016). Based on the framework of

Brattico et al., (2013), we assume AE is made up of perceptual, cognitive, affective, and aesthetic responses (e.g., liking). These components can be relatively well distinguished by self-reports and – to an extent – by different brain regions and event-related brain potentials, depending on their latency and polarity (e.g., early components are related to early sensory processes). However, changes in physiology/facial muscle activity have been related to all of these cognitive, affective, and aesthetic responses (e.g., Roy et al., 2009; Salimpoor et al., 2009; Steinbeis et al., 2006), depending on the design and control condition of the study. Some show physiological changes related to sensory (orienting response, e.g., Barry & Sokolov, 1993) and acoustic changes (e.g., Chuen et al., 2016), while others show this activity is related to aesthetic preference (e.g., Grewe et al., 2009; Salimpoor et al., 2009). In further understanding physiological responses, we draw on neural and behavioural evidence that gives better insight into what kind of AE-related processing might take place. On the one hand, responses related to sensory processing should be greater in the AO condition, due to less predictable sound onsets (Jessen & Kotz, 2011), as also shown by Vuoskoski et al. (2016). On the other hand, AV information conveys more emotion (Dahl & Friberg, 2007; Van Zijl & Luck, 2013); therefore, responses could also be higher in the AV condition, as shown in Chapados and Levitin (2008). Thus, we tested again whether physiological responses are higher in AO or AV.

We consistently found that the LF/HF ratio increased in the AO condition. As this measure represents increased SNS activation, this suggests that AO conditions increase physiological arousal, likely reflecting an increase in uncertainty of sound onsets when visual information is absent. This is in line with results from Vuoskoski et al. (2016), who found that AO evoked more physiological arousal (as shown by skin conductance) compared to AV musical performances. We also support findings by Richardson et al. (2020) who likewise found higher physiological arousal in audio-only, compared to video versions of narratives (e.g., *Games of Thrones* and *Pride and Prejudice*).

We also found partial support for the hypothesis that AV music performances lead to higher peripheral physiological responses than in AO performances. We state partial evidence, as design-driven LMMs differed from error-free ones. Simplified, error-free models revealed a significant modality effect for RR and EMGCS. Maximal models, which generated errors, did not. These differences could be attributed to the fact that removing the slopes to avoid singularity fit errors could have increased degrees of freedom and the possibility of Type 1 errors (Arnqvist, 2019). However, a model with a complex random-effects structure can lead to increased Type II error and lack of power (Barr, 2021; Matuschek et al., 2017). Thus, future studies with larger sample sizes are required to confirm this modality effect. As there is general consensus that error-free models are preferable (Barr et al., 2013), these models are reported. Nonetheless, we aim to be transparent; the reader is pointed to not only the Supplementary Materials, but also the code showing the maximal models and how models are simplified step by step. While only cautiously interpreting the modality effects in RR and EMGCS, we believe it is worth briefly discussing the results from error-free models.

RR was faster in the AV condition. 'Frowning' muscle (EMGCS) activity, which typically reflects negative valence (Bradley & Lang, 2000), also increased in the AV condition. The discrepancy between the increase in both frowning muscle activity and (generally positive) AE in the AV condition could be explained by the fact that higher aesthetic pleasure can also derive from perceiving negatively valenced musical expression and/or affective states (Eerola et al., 2018), such as being moved (Eerola et al., 2016). However, some question whether facial expressions reflect valence (Wingenbach et al., 2020) or affective states at all (Lewis, 2011; Matsumoto, 1987). Thus, another possible interpretation is that observing the musician increased mimicry in the observers. Indeed, participants mimic observed facial expressions (Dimberg, 1982; Magnee et al., 2007). Additionally, viewing effortful movements increases respiration (Brown et al., 2013; Mulder et al., 2005; Paccalin & Jeannerod, 2000) and corrugator activity (de Morree & Marcora, 2010). Such motor mimicry likely extends to music performance. Motor activity increases when listening to music (Bangert et al., 2006; Grahn & Brett, 2007; Janata et al., 2012), especially in audiovisual performances (Chan et al., 2013; Griffiths & Reay, 2018). Indeed, sensorimotor embodied mechanisms related to motor mimicry have been proposed and shown to enhance AE (Brattico & Pearce, 2013; Cross, 2011; Freedberg & Gallese, 2007; Gallese & Freedberg, 2007). Thus, faster breathing and increased facial muscle activity in AV conditions may be a reflection of motor mimicry that occurs when viewing musicians' movements. In sum, we provide partial evidence of a modality effect in RR and EMGCS, potentially reflecting motor mimicry.

Facial muscle activity was significantly associated with AE. The zygomaticus ('smiling') muscle activity was a significant predictor for AE scores. Increased zygomaticus activity was positively related to AE, supporting previous work showing that zygomaticus activity is higher for pleasant music (Fuentes-Sánchez et al., 2022), liked positive music (Witvliet & Vrana, 2007), positively evaluated art (Gernot et al., 2018), and liked dance movements (Kirsch et al., 2016). This is further support for the embodied aesthetics theory, where sensorimotor embodied mechanisms might enhance AE (Brattico & Pearce, 2013; Cross, 2011; Freedberg & Gallese, 2007; Gallese & Freedberg, 2007). However, increased 'smiling' muscle activity has also been shown to increase in unpleasant (dissonant) music, suggesting that such activity might represent a grimace or ironic laughter (Dellacherie et al., 2011; Merrill et al., 2021). Therefore, it is vital to collect self-report data to support interpretations of physiological responses, rather than considering certain responses a direct index of a specific state, especially over a long period of time in such naturalistic settings.

LMMs show that LF/HF ratio were higher in AO, and tentative evidence suggests that respiration and muscle activity were higher in AV. These findings can be considered in conjunction with how much (in)voluntary control we have over them. As mentioned before, the heart is innervated by the ANS and made up of involuntary (cardiac) muscle. Voluntary skeletal muscles control EMG and (partly) respiration. On the one hand, due to the automatic nature of the heart, it seems plausible these might be more related to earlier (sensory) processes of an AE. On the other hand, the more voluntary peripheral measures seem to be related to the liking aspect of AE. Although we are cautious to attribute the increase of such measures as a direct index of aesthetic experience, the results point to the idea that the more voluntary the control of the peripheral measure, the more related it may be to later stages of the aesthetic processing, as outlined in Brattico and Pearce (2013).

One overall limitation of the current study is that although all versions were presented as part of a concert while participants were seated in the concert hall, AV was presented as a live version, while AO was presented as a playback. This was chosen to enhance ecological validity: people who listen to music in an AO version most likely listen to music as playback, while watching an AV version is more likely to be live (Sloboda et al., 2012). Indeed, this difference of visual information is also shown in Swarbrick et al. (2019), who similarly stated that AV performances are typically live. Although we do appreciate that tools and streaming platforms like YouTube, Digital Concert Hall of the Berliner Philharmoniker and MetOnDemand etc. have increased in popularity (especially with the COVID-19 pandemic) making audio-visual recording more popular, Belfi et al. (2021) found that felt pleasure did not differ between live and an audio-visual recording of that same performance. Therefore, it is likely that the live and playback differences do not play a strong role in influencing the current results. Future research might consider live audio-only playback of an offstage performer to fully mitigate this potential confound. Another limitation is that although the pieces were chosen to represent typical concert pieces (and a range of genres), they were not controlled for length. Nonetheless, length was a compromise when using naturalistic stimuli that heightened ecological validity. As we did not look at piece-specific differences, but rather average across sections of the pieces to examine the effect of condition, we did not consider this a confound in the current study. However, we note that effects driven by one piece may weigh our results more heavily than effects from the shorter pieces. Future research might consider choosing pieces of similar length, or at least similar lengths of sections. A further limitation is that we did not contrast visual only information with the other two conditions. This choice was a compromise to keep the within-in subject design time-manageable as well as to create a concert-like feel for the experiment.

6 Conclusion

Researchers are increasingly foregoing ultimate control for a more ecologically valid approach that enables participants to have more powerful aesthetic experiences. This study follows others that have moved more into the 'wild' to explore such naturalistic experiences (Chabin et al., 2022; Czepiel et al., 2021; Dotov & Trainor, 2021; Merrill et al., 2021; Swarbrick et al., 2019; Tervaniemi et al., 2021). The current findings show that a self-reported aesthetic experience significantly increases in audio-visual (compared to audio-only) piano performances in the naturalistic setting of a concert hall.

Modality additionally influenced peripheral measures, revealing two main patterns. On the one hand, involuntary a physiological arousal response (heart rhythm reflecting SNS), was higher in the (less predictive) AO modality, likely reflecting more sensory processes. On the other hand, peripheral responses with more voluntary control (respiration, facial muscle activity) were higher in the AV modality, though due to inconsistencies in maximal/error-free models, these results should be interpreted with caution. The zygomaticus muscle was a significant predictor of self-reported AE. It could be that the involuntary-voluntary continuum of physiological responses is related to a sensory-affective continuum of AEs. We also suggest that visual information enhances motor mimicry (as shown by an increase in respiration and facial muscle activity), which is a mechanism that enhances AE (Cross et al., 2011; Freedberg & Gallese, 2007; Gallese & Freedberg, 2007; Kirsch et al., 2016). By exploring modality effects, we postulate that peripheral responses likely reflect sensory, sensorimotor, and affective responses that may culminate into an overall aesthetic experience (Brattico et al., 2013). However, we would like to emphasise that such peripheral responses alone cannot directly index AE; self-reports should support interpretations of peripheral physiological data. Nonetheless, the extent that physiological responses are simply sensory or reflect intertwined sensory and affective aspects of the aesthetic experience remains unclear. Further research, with larger sample sizes, should assess the robustness of the effects discussed here. To gain more insight, future research could bridge this gap by further exploring whether this involuntary-voluntary continuum reflects such sensory-aesthetic continuum and whether - and to what extent - there is an overlap of such systems.

Acknowledgements

The authors thank Lea T. Fink for helping with music theoretical analysis, as well as the ArtLab team and assistants for help for the concerts. We thank Klaus Frieler for advice on statistical analysis and the Music Department of MPI EA in Frankfurt and Melanie Wald-Fuhrmann for discussions about the selection of musical pieces. We also thank the anonymous reviewers who provided insightful feedback for improving the manuscript.

Author contribution:

Conceptualisation: CS. Methodology: CS. Investigation: CS, MS. Formal analysis: AC, SAK, MS, LKF. Writing - Original draft: AC. Visualisation: AC, LKF. Writing - review and editing: LKF, SAK, MS, CS

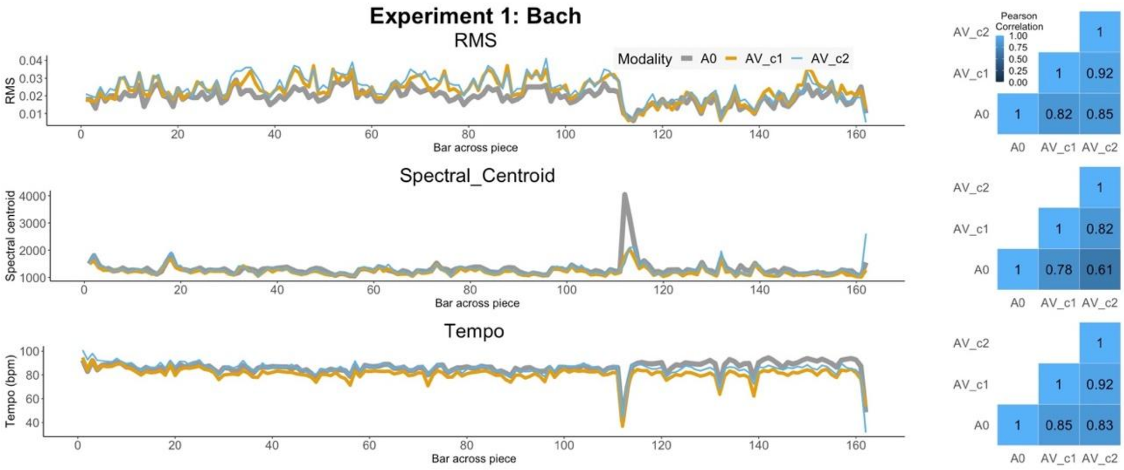
7 Supplementary Materials

Supplementary Table 1. Correlations of musical features between concerts.

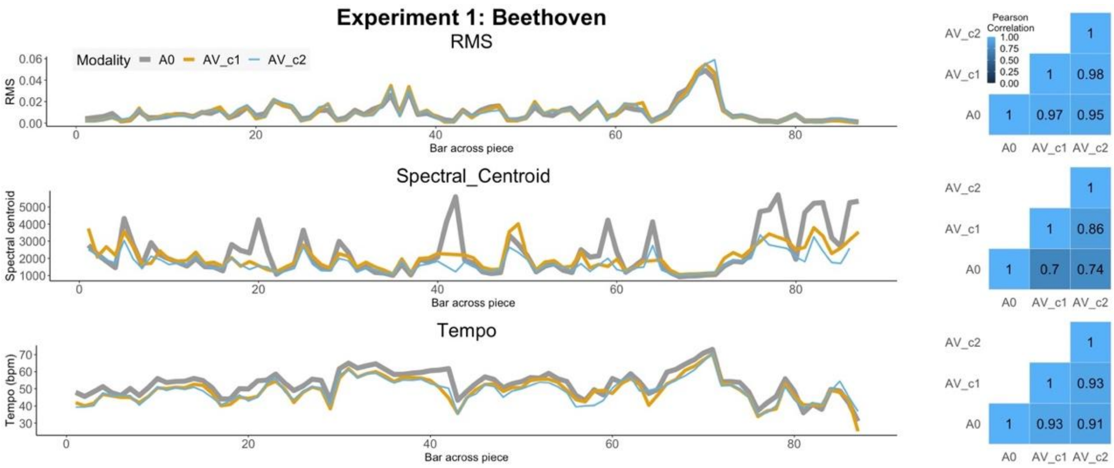
Experiment 1				
Feature	Piece	R: C1-C2 (*p*)	R: C1-AO (*p*)	R: C2-AO (*p*)
RMS	Bach	0.9185 ($p < .001$)	0.8229 ($p < .001$)	0.8539 ($p < .001$)
	Beet	0.9777 ($p < .001$)	0.9738 ($p < .001$)	0.9537 ($p < .001$)
	Mess	0.8813 ($p < .001$)	0.8028 ($p < .001$)	0.8196 ($p < .001$)
Spectral Centroid	Bach	0.8160 ($p < .001$)	0.7845 ($p < .001$)	0.6093 ($p < .001$)
	Beet	0.8568 ($p < .001$)	0.7033 ($p < .001$)	0.7400 ($p < .001$)
	Mess	0.8869 ($p < .001$)	0.8127 ($p < .001$)	0.7751 ($p < .001$)
Tempo	Bach	0.9195 ($p < .001$)	0.8496 ($p < .001$)	0.8316 ($p < .001$)
	Beet	0.9328 ($p < .001$)	0.9292 ($p < .001$)	0.9136 ($p < .001$)
	Mess	0.8861 ($p < .001$)	0.8883 ($p < .001$)	0.9283 ($p < .001$)
Experiment 2				
Feature	Piece	R: C3-C4 (*p*)	R: C3-AO (*p*)	R: C4-AO (*p*)
RMS	Bach	0.9498 ($p < .001$)	0.8976 ($p < .001$)	0.8969 ($p < .001$)
	Beet	0.9898 ($p < .001$)	0.9793($p < .001$)	0.9844 ($p < .001$)
	Mess	0.9574 ($p < .001$)	0.9211 ($p < .001$)	0.9194 ($p < .001$)
Spectral Centroid	Bach	0.9539 ($p < .001$)	0.6171 ($p < .001$)	0.6415 ($p < .001$)
	Beet	0.8931 ($p < .001$)	0.8512 ($p < .001$)	0.8200 ($p < .001$)
	Mess	0.9222 ($p < .001$)	0.9038 ($p < .001$)	0.8805 ($p < .001$)
Tempo	Bach	0.9337 ($p < .001$)	0.9495 ($p < .001$)	0.9583 ($p < .001$)
	Beet	0.9713 ($p < .001$)	0.9590 ($p < .001$)	0.9706 ($p < .001$)
	Mess	0.9815 ($p < .001$)	0.9597 ($p < .001$)	0.9706 ($p < .001$)

Supplementary Table 2. Information and bar numbers of sections that pieces were divided into, driven by the musical structure

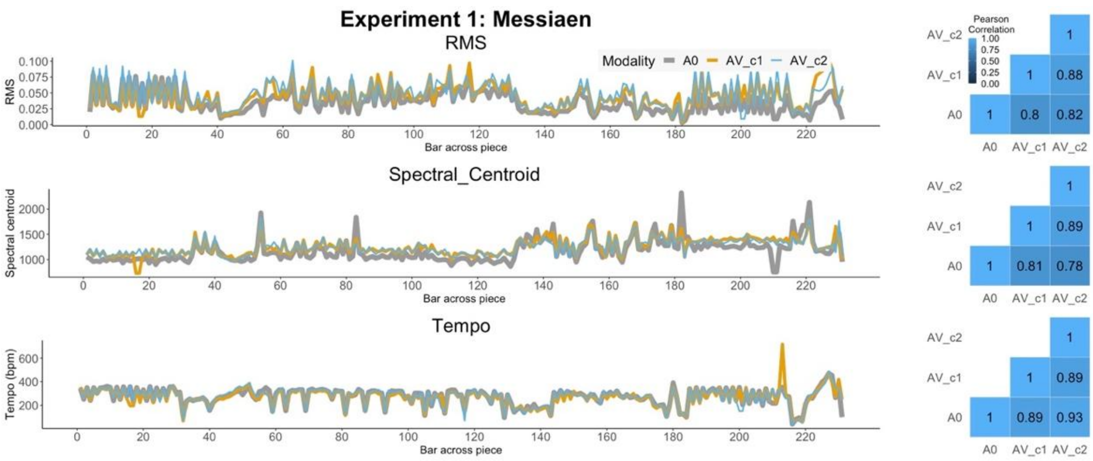
Piece	Section	Bar corresponding to music	Bars corresponding to acoustic and physiological signal (i.e., considering repeats)	Approximate length in time (seconds)
Bach	1	Prelude: 1-16	1-16	44
	2	Prelude: Repeat of bars 1-16	17-32	45
	3	Prelude: 17-40	33-56	68
	4	Prelude: 41 - 56	57-72	45
	5	Prelude: Repeat of bars 17-40	73-96	68
	6	Prelude: Repeat of bars 41-56	97-112	51
	7	Fugue: 1-50	113-162	137
Beethoven	1	1-8	1-8	60
	2	9-16	9-16	46
	3	17-25	17-25	64
	4	26-29	26-29	27
	5	30-43	30-43	42
	6	44-55	44-55	42
	7	56-64	56-64	81
	8	65-75	65-63	64
	9	76-87	76-87	68
Messiaen	1	1-32	1-32	60
	2	33-40	33-40	41
	3	41-59	41-59	37
	4	60-131	60-131	142
	5	132-143	132-143	68
	6	144-174	144-174	43
	7	175-184	175-184	55
	8	185-216	185-216	64
	9	217-231	217-231	49



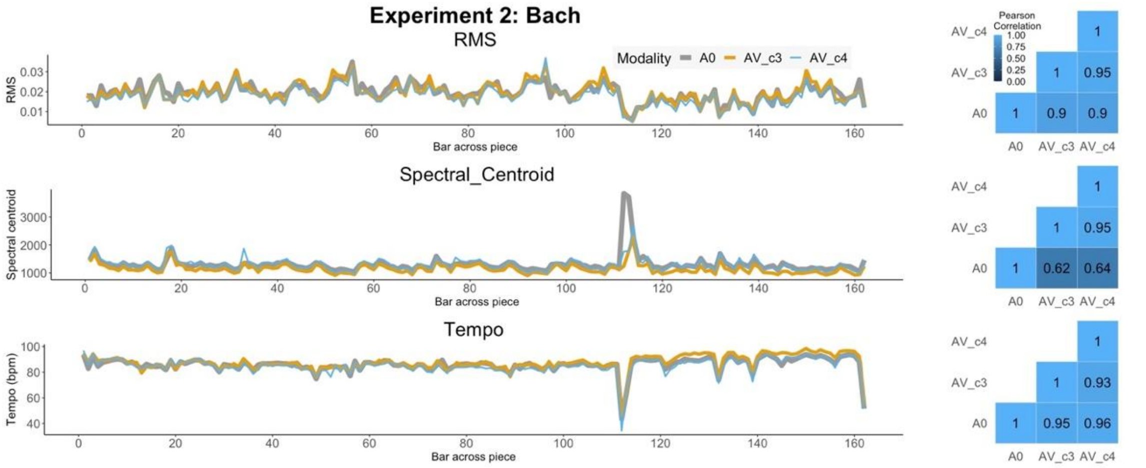
Supplementary Figure 1. Experiment 1, Bach piece: time series (left panels) and correlation matrices (right panels) of RMS, spectral centroid, and tempo (different rows) in AO and AV performances from concert 1 and 2.



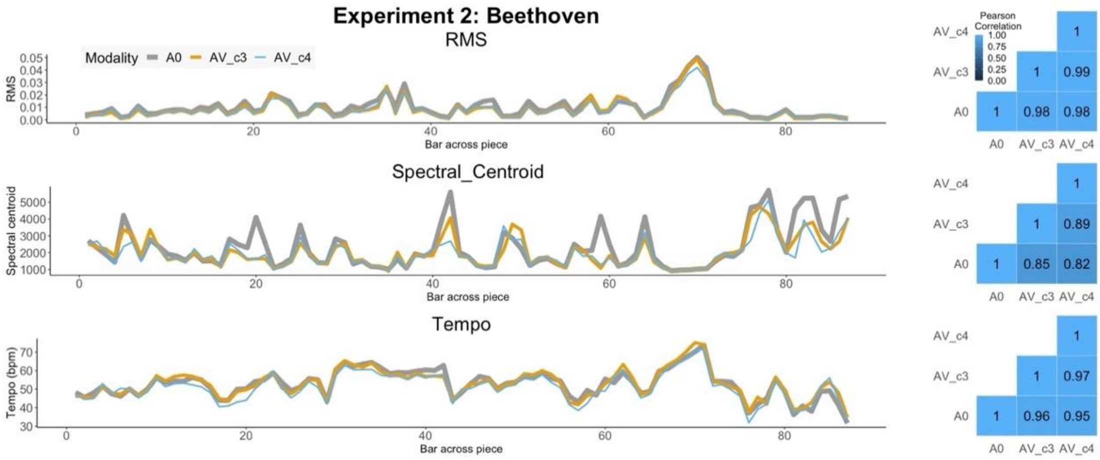
Supplementary Figure 2. Experiment 1, Beethoven piece: time series (left panels) and correlation matrices (right panels) of RMS, spectral centroid, and tempo (different rows) in AO and AV performances from concert 1 and 2.



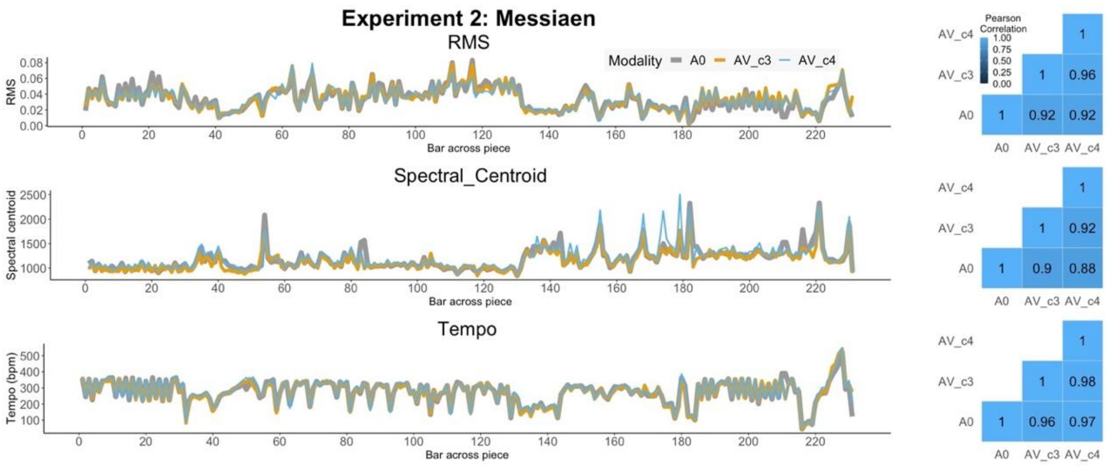
Supplementary Figure 3. Experiment 1, Messiaen piece: time series (left panels) and correlation matrices (right panels) of RMS, spectral centroid, and tempo (different rows) in AO and AV performances from concert 1 and 2.



Supplementary Figure 4. Experiment 2, Bach piece: time series (left panels) and correlation matrices (right panels) of RMS, spectral centroid, and tempo (different rows) in AO and AV performances from concert 3 and 4.



Supplementary Figure 5. Experiment 2, Beethoven piece: time series (left panels) and correlation matrices (right panels) of RMS, spectral centroid, and tempo (different rows) in AO and AV performances from concert 3 and 4.



Supplementary Figure 6. Experiment 2, Messiaen piece: time series (left panels) and correlation matrices (right panels) of RMS, spectral centroid, and tempo (different rows) in AO and AV performances from concert 3 and 4.

Supplementary Table 3. Linear mixed models for aesthetic experience factor in Experiment 1: maximal and simplified random structure models, until no error is generated.

LMM comparison for AE: Experiment 1

	Maximal model (singularity error)			**Constrained covariance parameters (singularity error)**			**Without slope (singularity error)**		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
(Intercept)	-0.10	-0.69 – 0.49	0.733	-0.10	-0.69 – 0.49	0.733	-0.10	-0.69 – 0.48	0.733
cond [AV]	0.29	0.05 – 0.52	**0.018**	0.29	0.05 – 0.52	**0.018**	0.29	0.05 – 0.52	**0.018**
Random Effects									
σ^2	0.55			0.55			0.55		
τ_{00}	$0.04_{\text{id\_n:concert}}$			$0.04_{\text{id\_n:concert}}$			$0.04_{\text{id\_n:concert}}$		
	0.24_{piece}			0.24_{piece}			0.24_{piece}		
	0.00_{concert}			0.00_{concert}			0.00_{concert}		
τ_{11}	$0.00_{\text{id\_n.condAV}}$			$0.00_{\text{id\_n.condAV}}$					
	$0.00_{\text{id\_n1.condAO}}$			$0.00_{\text{id\_n1.condAO}}$					
	$0.00_{\text{id\_n2.condAV}}$			$0.00_{\text{id\_n2.condAV}}$					
ρ_{01}									
ρ_{01}									
N	2_{concert}			2_{concert}			2_{concert}		
	3_{piece}			3_{piece}			3_{piece}		
	$15_{\text{id\_n}}$			$15_{\text{id\_n}}$			$15_{\text{id\_n}}$		
Observations	153			153			153		
Marginal R^2 / Conditional R^2	0.037 / NA			0.037 / NA			0.037 / NA		
AIC	379.967			379.967			373.967		

LMM comparison for AE: Experiment 1 (continued)

	Without concert intercept (singularity error)			**Without slope nor concert intercept (no error)**		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
(Intercept)	-0.10	-0.69 – 0.48	0.733	-0.10	-0.69 – 0.48	0.733
cond [AV]	0.29	0.05 – 0.52	**0.018**	0.29	0.05 – 0.52	**0.018**
Random Effects						
σ^2	0.55			0.55		
τ_{00}	$0.04_{\text{id\_n:concert}}$			$0.04_{\text{id\_n:concert}}$		
	0.24_{piece}			0.24_{piece}		
τ_{11}	$0.00_{\text{id\_n.condAV}}$					
	$0.00_{\text{id\_n1.condAO}}$					
	$0.00_{\text{id\_n2.condAV}}$					
ρ_{01}						
ρ_{01}						
ICC				0.34		
N	3_{piece}			3_{piece}		
	$15_{\text{id\_n}}$			$15_{\text{id\_n}}$		
	2_{concert}			2_{concert}		
Observations	153			153		
Marginal R^2 / Conditional R^2	0.037 / NA			0.024 / 0.355		
AIC	377.967			371.967		

Supplementary Table 4. Linear mixed models for aesthetic experience factor in Experiment 2: maximal and simplified random structure models, until no error is generated.

LMM comparison for AE: Experiment 2

	Maximal model (singularity error)			**Constrained covariance parameters (singularity error)**			**Without slope (singularity error)**		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
(Intercept)	0.00	-0.45 – 0.46	0.987	0.00	-0.45 – 0.46	0.987	0.00	-0.45 – 0.46	0.987
cond [AV]	0.22	0.01 – 0.43	**0.040**	0.22	0.01 – 0.43	**0.040**	0.22	0.01 – 0.43	**0.040**
Random Effects									
σ^2	0.40			0.40			0.40		
τ_{00}	0.15 id_n:concert			0.15 id_n:concert			0.16 id_n:concert		
	0.00 piece			0.00 piece			0.00 piece		
	0.08 concert			0.08 concert			0.08 concert		
τ_{11}	0.00 id_n1.condAO			0.00 id_n1.condAO					
	0.00 id_n2.condAV			0.00 id_n2.condAV					
ρ_{01}									
ρ_{01}									
N	2 concert			2 concert			2 concert		
	3 piece			3 piece			3 piece		
	16 id_n			16 id_n			16 id_n		
Observations	145			145			145		
Marginal R^2 / Conditional R^2	0.029 / NA			0.029 / NA			0.029 / NA		
AIC	332.563			332.563			326.565		

LMM comparison for AE: Experiment 2 (continued)

	Without piece intercept (singularity error)			**Without slope nor piece intercept (no error)**		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
(Intercept)	0.00	-0.45 – 0.46	0.987	0.00	-0.45 – 0.46	0.987
cond [AV]	0.22	0.01 – 0.43	**0.040**	0.22	0.01 – 0.43	**0.040**
Random Effects						
σ^2	0.40			0.40		
τ_{00}	0.16 id_n:concert			0.16 id_n:concert		
	0.08 concert			0.08 concert		
τ_{11}	0.00 id_n.condAV					
	0.00 id_n1.condAO					
	0.00 id_n2.condAV					
ρ_{01}						
ρ_{01}						
ICC				0.37		
N	2 concert			2 concert		
	16 id_n			16 id_n		
Observations	145			145		
Marginal R^2 / Conditional R^2	0.029 / NA			0.019 / 0.383		
AIC	330.565			324.565		

Supplementary Table 5. Linear mixed models for EMGCS (Corrugator facial muscle activity) in Experiment 2: maximal and simplified random structure models, until no error is generated.

LMM comparison for EMGCS: Experiment 2

	Maximal model (singularity error)			Constrained covariance parameters (error)			Without slope (error)		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
(Intercept)	0.0025	0.0020 – 0.0030	**<0.001**	0.0025	0.0020 – 0.0030	**<0.001**	0.0025	0.0021 – 0.0029	**<0.001**
cond [AV]	0.0002	-0.0000 – 0.0004	0.098	0.0002	-0.0000 – 0.0004	0.098	0.0002	0.0001 – 0.0003	**<0.001**
Random Effects									
σ^2	0.00			0.00			0.00		
τ_{00}	0.00 section:piece			0.00 section:piece			0.00 section:piece		
	0.00 id_n:concert			0.00 id_n:concert			0.00 id_n:concert		
	0.00 concert			0.00 concert			0.00 concert		
τ_{11}	0.00 id_n.condAV			0.00 id_n.condAV					
	0.00 id_n1.condAO			0.00 id_n1.condAO					
	0.00 id_n2.condAV			0.00 id_n2.condAV					
ρ_{01}									
ρ_{01}									
ICC	0.61			0.61			0.65		
N	2 concert			2 concert			2 concert		
	9 section			9 section			9 section		
	3 piece			3 piece			3 piece		
	15 id_n			15 id_n			15 id_n		
Observations	1037			1037			1037		
Marginal R^2 / Conditional R^2	0.010 / 0.617			0.010 / 0.617			0.007 / 0.657		
AIC	-12173.674			-12173.674			-12108.532		

LMM comparison for EMGCS: Experiment 2 (continued)

	Without concert intercept (error)			Without slope nor concert intercept (no error)		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
(Intercept)	0.0025	0.0020 – 0.0030	**<0.001**	0.0025	0.0021 – 0.0029	**<0.001**
cond [AV]	0.0002	-0.0000 – 0.0004	0.098	0.0002	0.0001 – 0.0003	**<0.001**
Random Effects						
σ^2	0.00			0.00		
τ_{00}	0.00 section:piece			0.00 section:piece		
	0.00 id_n:concert			0.00 id_n:concert		
τ_{11}	0.00 id_n.condAV					
	0.00 id_n1.condAO					
	0.00 id_n2.condAV					
ρ_{01}						
ρ_{01}						
ICC	0.61			0.65		
N	9 section			9 section		
	3 piece			3 piece		
	15 id_n			15 id_n		
	2 concert			2 concert		
Observations	1037			1037		
Marginal R^2 / Conditional R^2	0.010 / 0.618			0.007 / 0.657		
AIC	-12175.674			-12110.532		

Supplementary Table 6. Linear mixed models for EMGZM (Zygomaticus facial muscle activity) in Experiment 2: maximal and simplified random structure models, until no error is generated.

LMM comparison for EMGZM: Experiment 2

	Maximal model (singularity error)			Constrained covariance parameters (error)		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
(Intercept)	0.0022	0.0019 – 0.0024	**<0.001**	0.0022	0.0019 – 0.0024	**<0.001**
cond [AV]	0.0001	-0.0000 – 0.0002	0.180	0.0001	-0.0000 – 0.0002	0.180
Random Effects						
σ^2	0.00			0.00		
τ_{00}	0.00 section:piece			0.00 section:piece		
	0.00 id_n:concert			0.00 id_n:concert		
	0.00 concert			0.00 concert		
τ_{11}	0.00 id_n.condAV			0.00 id_n.condAV		
	0.00 id_n1.condAO			0.00 id_n1.condAO		
	0.00 id_n2.condAV			0.00 id_n2.condAV		
ρ_{01}						
ρ_{01}						
N	2 concert			2 concert		
	9 section			9 section		
	3 piece			3 piece		
	14 id_n			14 id_n		
Observations	1082			1082		
Marginal R^2 / Conditional R^2	0.006 / NA			0.006 / NA		
AIC	-12780.993			-12780.993		

LMM comparison for EMGZM: Experiment 2 (continued)

	Without slope (error)			Without concert intercept (no error)		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
(Intercept)	0.0022	0.0019 – 0.0024	**<0.001**	0.0022	0.0019 – 0.0024	**<0.001**
cond [AV]	0.0001	-0.0000 – 0.0001	0.052	0.0001	-0.0000 – 0.0002	0.180
Random Effects						
σ^2	0.00			0.00		
τ_{00}	0.00 section:piece			0.00 section:piece		
	0.00 id_n:concert			0.00 id_n:concert		
	0.00 concert					
τ_{11}				0.00 id_n.condAV		
				0.00 id_n1.condAO		
				0.00 id_n2.condAV		
ρ_{01}						
ρ_{01}						
ICC	0.50			0.50		
N	2 concert			9 section		
	9 section			3 piece		
	3 piece			14 id_n		
	14 id_n			2 concert		
Observations	1082			1082		
Marginal R^2 / Conditional R^2	0.002 / 0.506			0.003 / 0.502		
AIC	-12771.767			-12782.993		

Supplementary Table 7. Linear mixed models for High Frequency power in heart rate variability (HF) in Experiment 2: maximal and simplified random structure models, until no error is generated.

LMM comparison for HF: Experiment 2

	Maximal model (singularity error)			Constrained covariance parameters (error)		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
(Intercept)	0.1384	0.1147 – 0.1621	**<0.001**	0.1384	0.1147 – 0.1621	**<0.001**
cond [AV]	0.0044	-0.0069 – 0.0157	0.447	0.0044	-0.0069 – 0.0157	0.447

Random Effects

σ^2	0.00	0.00
τ_{00}	0.00 section:piece	0.00 section:piece
	0.00 id_n:concert	0.00 id_n:concert
	0.00 concert	0.00 concert
τ_{11}	0.00 id_n1.condAO	0.00 id_n1.condAO
	0.00 id_n2.condAV	0.00 id_n2.condAV
ρ_{01}		
ρ_{01}		
ICC	0.47	0.47
N	2 concert	2 concert
	9 section	9 section
	3 piece	3 piece
	15 id_n	15 id_n
Observations	1050	1050
Marginal R^2 / Conditional R^2	0.001 / 0.467	0.001 / 0.467
AIC	-2839.267	-2839.267

LMM comparison for HF: Experiment 2 (continued)

	Without slope (error)			Without concert intercept (no error)		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
(Intercept)	0.1394	0.1160 – 0.1628	**<0.001**	0.1384	0.1147 – 0.1621	**<0.001**
cond [AV]	0.0068	-0.0004 – 0.0139	0.066	0.0044	-0.0069 – 0.0157	0.447

Random Effects

σ^2	0.00	0.00
τ_{00}	0.00 section:piece	0.00 section:piece
	0.00 id_n:concert	0.00 id_n:concert
	0.00 concert	
τ_{11}		0.00 id_n1.condAO
		0.00 id_n2.condAV
ρ_{01}		
ρ_{01}		
ICC		0.47
N	2 concert	9 section
	9 section	3 piece
	3 piece	15 id_n
	15 id_n	2 concert
Observations	1050	1050
Marginal R^2 / Conditional R^2	0.003 / NA	0.001 / 0.467
AIC	-2837.652	-2841.267

Supplementary Table 8. Linear mixed models for Low Frequency / High Frequency ratio (LF/HF ratio) in heart rate variability in Experiment 2: maximal and simplified random structure models, until no error is generated.

LMM comparison for LFHF ratio: Experiment 2

	Maximal model (singularity error)			Constrained covariance parameters (error)			Without slope (error)		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
(Intercept)	2.1904	1.7790 – 2.6017	**<0.001**	2.1904	1.7790 – 2.6017	**<0.001**	2.1926	1.7794 – 2.6059	**<0.001**
cond [AV]	-0.2396	-0.4304 – -0.0488	**0.014**	-0.2396	-0.4304 – -0.0488	**0.014**	-0.2604	-0.4116 – -0.1093	**0.001**
Random Effects									
σ^2	1.51			1.51			1.52		
τ_{00}	0.00 section:piece			0.00 section:piece			0.00 section:piece		
	0.94 id_n:concert			0.94 id_n:concert			0.95 id_n:concert		
	0.00 concert			0.00 concert			0.00 concert		
τ_{11}	0.00 id_n1.condAO			0.00 id_n1.condAO					
	0.04 id_n2.condAV			0.04 id_n2.condAV					
ρ_{01}									
ρ_{01}									
ICC	0.39			0.39			0.38		
N	2 concert			2 concert			2 concert		
	9 section			9 section			9 section		
	3 piece			3 piece			3 piece		
	15 id_n			15 id_n			15 id_n		
Observations	1041			1041			1041		
Marginal R^2 / Conditional R^2	0.006 / 0.392			0.006 / 0.392			0.007 / 0.389		
AIC	3489.670			3489.670			3485.161		

LMM comparison for LFHF ratio: Experiment 2 (continued)

	Without concert intercept (error)			Without slope nor concert intercept (no error)		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
(Intercept)	2.1903	1.7792 – 2.6015	**<0.001**	2.1926	1.7794 – 2.6059	**<0.001**
cond [AV]	-0.2396	-0.4304 – -0.0488	**0.014**	-0.2604	-0.4116 – -0.1093	**0.001**
Random Effects						
σ^2	1.51			1.52		
τ_{00}	0.00 section:piece			0.00 section:piece		
	0.94 id_n:concert			0.95 id_n:concert		
τ_{11}	0.00 id_n1.condAO					
	0.04 id_n2.condAV					
ρ_{01}						
ρ_{01}						
ICC	0.39			0.38		
N	9 section			9 section		
	3 piece			3 piece		
	15 id_n			15 id_n		
	2 concert			2 concert		
Observations	1041			1041		
Marginal R^2 / Conditional R^2	0.006 / 0.392			0.007 / 0.389		
AIC	3487.670			3483.161		

Supplementary Table 9. Linear mixed models for heart rate (HR) in Experiment 2: maximal and simplified random structure models, until no error is generated.

LMM comparison for HR: Experiment 2

	Maximal model (error)			Constrained covariance parameters (error)			Without slope (no error)		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
(Intercept)	62.0143	54.3396 – 69.6890	**<0.001**	62.0143	54.3396 – 69.6890	**<0.001**	61.9969	54.9576 – 69.0361	**<0.001**
cond [AV]	-0.2951	-1.2997 – 0.7094	0.564	-0.2951	-1.2997 – 0.7094	0.564	-0.1932	-0.4390 – 0.0527	0.123

Random Effects

	Maximal model (error)	Constrained covariance parameters (error)	Without slope (no error)
σ^2	3.46	3.46	4.16
τ_{00}	0.13 section:piece	0.13 section:piece	0.11 section:piece
	65.05 id_n:concert	65.05 id_n:concert	65.70 id_n:concert
	24.91 concert	24.91 concert	19.99 concert
τ_{11}	0.00 id_n1.condAO	0.00 id_n1.condAO	
	3.72 id_n2.condAV	3.72 id_n2.condAV	
ρ_{01}			
ρ_{01}			
ICC			0.95
N	2 concert	2 concert	2 concert
	9 section	9 section	9 section
	3 piece	3 piece	3 piece
	15 id_n	15 id_n	15 id_n
Observations	1073	1073	1073
Marginal R^2 / Conditional R^2	0.006 / NA	0.006 / NA	0.000 / 0.954
AIC	4613.406	4613.406	4754.982

Supplementary Table 10. Linear mixed models for respiration rate (RR) in Experiment 2: maximal and simplified random structure models, until no error is generated.

LMM comparison for RR: Experiment 2

	Maximal model (singularity error)			Constrained covariance parameters (error)			Without slope (no error)		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
(Intercept)	18.2917	17.0021 – 19.5814	**<0.001**	18.2917	17.0021 – 19.5814	**<0.001**	18.3133	17.1729 – 19.4537	**<0.001**
cond [AV]	0.2274	-0.1377 – 0.5925	0.222	0.2274	-0.1377 – 0.5925	0.222	0.2570	0.0873 – 0.4266	**0.003**

Random Effects

	Maximal model (singularity error)	Constrained covariance parameters (error)	Without slope (no error)
σ^2	2.03	2.03	2.11
τ_{00}	0.31 section:piece	0.31 section:piece	0.30 section:piece
	4.71 id_n:concert	4.71 id_n:concert	6.71 id_n:concert
	0.00 concert	0.00 concert	0.08 concert
τ_{11}	3.17 id_n1.condAO	3.17 id_n1.condAO	
	2.04 id_n2.condAV	2.04 id_n2.condAV	
ρ_{01}			
ρ_{01}			
ICC	0.75	0.75	0.77
N	2 concert	2 concert	2 concert
	9 section	9 section	9 section
	3 piece	3 piece	3 piece
	15 id_n	15 id_n	15 id_n
Observations	1152	1152	1152
Marginal R^2 / Conditional R^2	0.002 / 0.749	0.002 / 0.749	0.002 / 0.771
AIC	4294.608	4294.608	4312.895

Supplementary Table 11. Linear mixed models for skin conductance response (SCR) in Experiment 2: maximal and simplified random structure models, until no error is generated.

LMM comparison for SCR: Experiment 2

	Maximal model (singularity error)			Constrained covariance parameters (error)			Without slope (error)		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
(Intercept)	-0.0019	-0.0045 – 0.0007	0.156	-0.0019	-0.0045 – 0.0007	0.156	-0.0019	-0.0045 – 0.0007	0.156
cond [AV]	-0.0002	-0.0016 – 0.0012	0.754	-0.0002	-0.0016 – 0.0012	0.754	-0.0002	-0.0016 – 0.0012	0.754
Random Effects									
σ^2	0.00			0.00			0.00		
τ_{00}	0.00 section:piece			0.00 section:piece			0.00 section:piece		
	0.00 id_n:concert			0.00 id_n:concert			0.00 id_n:concert		
	0.00 concert			0.00 concert			0.00 concert		
τ_{11}	0.00 id_n.condAV			0.00 id_n.condAV					
	0.00 id_n1.condAO			0.00 id_n1.condAO					
	0.00 id_n2.condAV			0.00 id_n2.condAV					
ρ_{01}									
ρ_{01}									
N	2 concert			2 concert			2 concert		
	9 section			9 section			9 section		
	3 piece			3 piece			3 piece		
	14 id_n			14 id_n			14 id_n		
Observations	855			855			855		
Marginal R^2 / Conditional R^2	0.000 / NA			0.000 / NA			0.000 / NA		
AIC	-5275.751			-5275.751			-5281.751		

LMM comparison for SCR: Experiment 2 (continued)

	Without concert intercept (error)			Without slope nor concert intercept (no error)		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
(Intercept)	-0.0019	-0.0045 – 0.0007	0.156	-0.0019	-0.0045 – 0.0007	0.156
cond [AV]	-0.0002	-0.0016 – 0.0012	0.754	-0.0002	-0.0016 – 0.0012	0.754
Random Effects						
σ^2	0.00			0.00		
τ_{00}	0.00 section:piece			0.00 section:piece		
	0.00 id_n:concert			0.00 id_n:concert		
τ_{11}	0.00 id_n.condAV					
	0.00 id_n1.condAO					
	0.00 id_n2.condAV					
ρ_{01}						
ρ_{01}						
ICC				0.25		
N	9 section			9 section		
	3 piece			3 piece		
	14 id_n			14 id_n		
	2 concert			2 concert		
Observations	855			855		
Marginal R^2 / Conditional R^2	0.000 / NA			0.000 / 0.252		
AIC	-5277.751			-5283.751		

Supplementary Table 12. Linear mixed models for skin conductance level (SCL) in Experiment 2: maximal and simplified random structure models, until no error is generated.

LMM comparison for SCL: Experiment 2

	Maximal model (singularity error)			Constrained covariance parameters (error)			Without slope (error)		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
(Intercept)	0.0005	-0.0247 – 0.0257	0.970	0.0005	-0.0247 – 0.0257	0.970	0.0005	-0.0247 – 0.0257	0.970
cond [AV]	0.0022	-0.0124 – 0.0168	0.764	0.0022	-0.0124 – 0.0168	0.764	0.0022	-0.0124 – 0.0168	0.764
Random Effects									
σ^2	0.01			0.01			0.01		
τ_{00}	0.00 section:piece			0.00 section:piece			0.00 section:piece		
	0.00 id_n:concert			0.00 id_n:concert			0.00 id_n:concert		
	0.00 concert			0.00 concert			0.00 concert		
τ_{11}	0.00 id_n.condAV			0.00 id_n.condAV					
	0.00 id_n1.condAO			0.00 id_n1.condAO					
	0.00 id_n2.condAV			0.00 id_n2.condAV					
ρ_{01}									
ρ_{01}									
N	2 concert			2 concert			2 concert		
	9 section			9 section			9 section		
	3 piece			3 piece			3 piece		
	14 id_n			14 id_n			14 id_n		
Observations	910			910			910		
Marginal R^2 / Conditional R^2	0.000 / NA			0.000 / NA			0.000 / NA		
AIC	-1312.587			-1312.587			-1318.587		

LMM comparison for SCL: Experiment 2 (continued)

	Without concert intercept (error)			Without slope nor concert intercept (error)			Without slope nor concert and id intercept (no error)		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
(Intercept)	0.0005	-0.0247 – 0.0257	0.970	0.0005	-0.0247 – 0.0257	0.970	0.0005	-0.0247 – 0.0257	0.970
cond [AV]	0.0022	-0.0124 – 0.0168	0.764	0.0022	-0.0124 – 0.0168	0.764	0.0022	-0.0124 – 0.0168	0.764
Random Effects									
σ^2	0.01			0.01			0.01		
τ_{00}	0.00 section:piece			0.00 section:piece			0.00 section:piece		
	0.00 id_n:concert			0.00 id_n:concert					
τ_{11}	0.00 id_n.condAV								
	0.00 id_n1.condAO								
	0.00 id_n2.condAV								
ρ_{01}									
ρ_{01}									
ICC							0.21		
N	9 section			9 section			9 section		
	3 piece			3 piece			3 piece		
	14 id_n			14 id_n					
	2 concert			2 concert					
Observations	910			910			910		
Marginal R^2 / Conditional R^2	0.000 / NA			0.000 / NA			0.000 / 0.213		
AIC	-1314.587			-1320.587			-1322.587		

Questionnaire items (translated from German into English)

0.1 Please state your age:

0.2 Please state your gender: 1. female; 2. male; 3. other

0.3 Please state your highest level of education:

1. Secondary school leaving certificate / Mittlere Reife
2. (Technical) Baccalaureate
3. Vocational training
4. College or university degree
5. not specified

0.4 How many years of instrumental lessons (including singing) have you had in your life? ______ years

0.5 Do you sing? (yes/no)

0.6 Do you play an instrument? (yes/no)

0.7 I would describe myself as a musician (1. do not agree to 7. Completely agree).

0.8 How many concerts/ live musical events have you attended within the last twelve months?

0.9 What kind of concerts/ live musical events do you attend most often? (rock/pop, classical, club/disco, jazz, contemporary, musical, opera, church, other).

1.1. Please answer the following questions

1.1.1 How much did you like the piece? (1-7)

1.1.2 How much did you like the interpretation? (1-7)

1.1.3 How familiar are you with this style of music that you have just heard? (1-7)

1.1.4 Do you know the piece? (yes/no)

1.2 To what extent do the following phrases apply to you?

1.2.1 I felt the need to move (1-7)

1.2.2 I tried to understand what was happening in the music (1-7)

1.2.3 I felt a connection to the musicians (1-7)

1.2.4 I was completely immersed in the music (1-7)

1.2.5 I felt connected to the other audience members (listeners) (1-7)

1.2.6 I was simply let the music affect me (passively receiving the music) (1-7)

1.2.7 I felt distracted by the measuring equipment (1-7)

1.3 Any other comments?

Chapter 5

Cardiorespiratory synchrony in audiences during music concerts

Based on:

Czepiel, A., Fink, L. K., Seibert, C., Scharinger M., Wald-Fuhrmann, M., Kotz, S. A. (In preparation). Cardiorespiratory synchrony in audiences during music concerts.

Abstract

People enjoy listening to and engaging with music, especially in live concerts. One key concert component that might enhance engagement is the visual aspect of seeing a musician perform. As recent evidence showed that neural and cardiorespiratory synchrony can index engagement with music, we tested the influence of modality in a concert setting, by comparing musical engagement – indexed by cardiorespiratory synchrony – in audio-only (AO) and audio-visual (AV) versions of musical performances. In a concert experiment, audiences were presented with AO and AV versions of piano performances, while cardiorespiratory measures were continuously recorded. Based on Kaneshiro et al. (2020), four synchrony measures were calculated to represent musical engagement: stimulus-response synchrony and inter-subject synchrony, both in the time and phase domains. After checking the significance of these synchrony measures by comparing true (time-locked) responses to control (circular-shifted) responses, we compared synchrony measures between AO and AV musical on a global level across musical pieces as well as on a time-resolved level at salient moments in the music, i.e., structural boundaries. Time domain synchrony measures were significantly higher than circularly shifted data. Musical engagement – as indexed by cardiorespiratory synchrony – was greater when audiences saw the musician, both at a global and time-resolved level. Intersubject correlation of HR was the most robust synchrony measure. This not only supports previous results, but also informs future studies as to which measures might be most informative to study musical engagement in ecological settings.

1 Introduction

People enjoy engaging with music, especially in live concerts (Gabrielsson & Wik, 2003; Lamont, 2011). One concert component that may enhance musical engagement – defined here as a listener's real-time absorption and interest with the music (Kaneshiro et al., 2020; Schubert et al., 2013) – is the visual aspect of seeing musicians perform (Wald-Fuhrmann et al., 2021). We have previously shown that audio-visual performances compared to audio-only performances increased aesthetic appreciation (Czepiel et al., 2021; Platz & Kopiez, 2012). These evaluations, however, were recorded at the end of musical pieces that were 7-12 minutes long. A further step would be to explore engagement and how it varies over time. Aiming to extend lab studies (Platz & Kopiez, 2012) to a setting that affords greater opportunity to engage with the music, the current study investigated dynamic engagement between audio-only and audio-visual music performances in a concert setting.

Although musical engagement can be measured via continuous behavioural responses, one's engagement with the stimuli might be decreased by the actual task of rating. Peripheral responses may be preferred as they do not distract from a music listening experience. In particular, cardiorespiratory measures are promising alternatives. For example, orienting and affective states in music listening have been respectively related to deceleration-acceleration patterns of heart rate (HR) (Bradley & Lang, 2000; Graham & Clifton, 1966; Stekelenburg & van Boxtel, 2002), and increased HR and respiration rate (RR) (Blood & Zatorre, 2001; Grewe et al., 2009; Guhn et al., 2007; Salimpoor et al., 2009). Of interest for the current paper, growing evidence demonstrates that engagement with the stimulus could be reflected by synchrony of neural and peripheral measures (e.g., Dmochowski et al., 2012; Pérez et al., 2021). The current study explores four kinds of cardiorespiratory synchrony in time and phase domains (Kaneshiro et al., 2020; Weineck et al., 2022).

In assessing the extent that musical stimuli can evoke time- and phase-locked cardiorespiratory responses (e.g., Egermann et al., 2013; Etzel et al., 2005; Gomez & Danuser, 2007; Haas et al. 1987; Sammler et al., 2007), synchrony between a stimulus and a participant's response was assessed, both in the time-domain (stimulus-response correlation, SRC or SRCorr) and phase-domain (Dmochowski et al., 2012; Kaneshiro et al., 2019; Weineck et al., 2022). For time-domain synchrony, this would refer to how much HR and RR are synchronised to a musical feature. For the phase domain, this would refer to whether heartbeats/ breaths align to rhythms of a stimulus. So far studies have shown that respiration entrains to musical beats (Etzel et al., 2006; Haas et al., 1986), but there is limited evidence that heart beats do the same (Ellis & Thayer, 2010; Koelsch & Jäncke, 2015; Mütze et al., 2020). A related approach to observe music-related responses is to assess synchrony across multiple participant responses that are time-locked to a stimulus. Again, this can either be in the time domain, referred to in the field as inter-subject correlation (ISC; Dmochowski et al., 2012; Hasson et al., 2004), or in the phase domain, exploring whether heart beats/breaths between participants align referred to here as inter-subject phase coherence (ISPC).

Stimulus-response and inter-subject synchrony both in the time and phase-domain have been related to engagement with auditory stimuli. In the time domain, SRC and ISC have been related to engagement with speech (Herrmann & Johnsrude, 2020; Irsik et al., 2022; Schmälzle et al., 2015), movies (Dmochowski et al., 2012; Kang & Wheatley, 2017; Ki et al., 2016), and music (Kaneshiro et al., 2020; Madsen et al., 2019). In the phase domain, interbrain phase synchrony has been found to be higher in shared engagement and engaging group discussions (Dikker et al., 2017, 2021). Although phase synchrony in auditory rhythms, such as speech and music, have been attributed to intelligibility to stimulus (Harding et al., 2019; Luo & Poeppel, 2007; Peelle et al., 2013; Vanden Bosch der Nederlanden et al., 2020) rather than engagement, it has nonetheless been postulated that phase synchrony increases when attending to stimuli. According to the Dynamic Attending Theory (DAT), internal oscillations adapt to external rhythms so that attending energy is optimised at expected points in time (Dynamic Attending Theory, DAT, Jones & Boltz, 1989; Large & Jones, 1999; Large & Snyder, 2009).

Most work on inter-subject and stimulus-response synchrony has been exclusively in neural responses such as EEG. Recent frameworks propose such neural mechanisms might extend to peripheral rhythms, cardiac, respiratory, gut and pupil dynamics (for reviews see Criscuolo et al., 2022; Klimesch et al., 2018; Parviainen, Lyyra, & Nokia, 2022). Indeed, some research indicates that the alignments of breathing and heart beats aid processing of upcoming events (Al et al., 2020; Criscuolo et al., 2022; Grund et al., 2022). There are already some promising results showing cardiorespiratory ISC to narratives and instructional videos (Madsen et al., 2022; Pérez et al., 2021, see also Palumbo et al., 2017). The current study aims to extend this work to assess cardiorespiratory synchrony in response to music performances.

Cardiorespiratory synchrony can be assessed on a global level to assess engagement, for example comparing synchrony across a whole stimulus when participants attended to, or were distracted from, the stimuli (Pérez et al., 2021). However, it is also possible to assess time-resolved synchrony at specific time points in the music (Czepiel et al., 2021; Dauer et al., 2021). For the current study, we wanted to not only assess synchrony measures between modality conditions overall, but also at specific time points. To this end, we assessed synchrony at epochs centred around salient moments in the music, that is, the perception of a distinct and prominently change in the music. Indeed, salient rhythmic, harmonic, and melodic changes all evoke increases in self-reported engagement (Presicce & Bailes, 2019; Schubert et al., 2013), peripheral responses (Chuen et al., 2016; Egermann et al., 2013; Fink et al., 2018; Schultz et al., 2021; Steinbeis et al., 2006) as well as increases in neural (Kaneshiro et al., 2016; Dauer et al., 2021) and peripheral synchrony (Czepiel et al., 2021). In obtaining optimal moments for such salient changes, we chose section boundaries as they are most likely to be perceived as salient moments (Lerdahl & Jackendoff, 1983). Thus, we examined the overall physiological synchrony, and as well as synchrony changes synchrony across epochs ±15 seconds relative to section boundaries.

The current study

It has been proposed that concert performances may be more engaging due to the visual aspect of seeing a musician. While this has been tested in a laboratory capacity, this study aimed to extend this work in a setting that affords greater capacity for engagement with music, namely a concert setting. With much work showing that engagement with a stimulus is related to stimulus-response and inter-subject synchrony of neural responses, this study aimed also to extend this to cardiorespiratory responses. To test the influence of modality on engagement to music, concert audiences were presented with audio-only (AO) and audio-visual (AV) presentations of natural piano pieces, with cardiorespiratory synchrony measures indexing dynamic engagement to music. Three hypotheses were tested. First, we tested the assumption that time-locked responses are significantly synchronised compared to circular-shifted control data. Second, that overall (global) synchrony would be higher in AV, than in AO musical performances. Third, we expected synchrony to be higher in AV at more localised, salient moments in the music, i.e., at structural boundaries. Observing a clear effect of structural boundary allows us to study how these typical responses differ according to modality.

2 Method

Participants, stimuli, and procedure are identical to Czepiel et al. (2023, bioRxiv, see General Method and Experiment 2). Key details of the procedure are outlined below.

Participants

The study was approved by the Ethics Council of the Max Planck Society and in accordance with the Declaration of Helsinki. Participants gave their written informed consent. Behavioural and physiological data of twenty-five participants were analysed (one participant was excluded due to missing physiological data). Participants consisted of nine females (16 males), mean age of 51.08 years (SD = 15.48), who on average had 6.18 (SD = 8.20) years of music lessons and attended on average 13 concerts per year (M = 19.97, SD = 20).

Stimuli

Three musical pieces were presented to the participants in AV and AO versions (see Figure 1). Both versions were performed by the same professional pianist, on the same piano, in the same concert hall. AO versions were recorded prior to the concert and were presented as recordings through high-quality loudspeakers (Fohhn LX-150 + Fohhn XS-22). AV versions were performed live by the pianist during the concert. A trained sound engineer checked that the sound level was equal across all stimuli and acoustic features analysis shows these performances were comparable across concerts.

Procedure

Participants were invited to attend one of two concerts which took place on consecutive evenings at a hybrid laboratory-concert hall (www.aesthetics.mpg.de/en/artlab/information.html). Electrocardiography (ECG) and respiration was collected using gelled self-adhesive disposable Ag/AgCl electrodes and a respiration belt, respectively. Prior to the concert starting, participants were debriefed and filled in an initial demographic questionnaire. During the concert, audiences heard the three pieces in both modalities. Order of modality was counterbalanced across concerts. After each piece, participants rated items such as absorption and understanding of the music. Physiological signals were continuously measured during the concert at a 1000 Hz sampling rate.

Data Analysis

Musical analysis: spectral flux and structural boundaries

To characterise the music, continuous spectral flux was obtained using the MIRToolbox in Matlab (Lartillot & Toiviainen, 2007), with a frame size of 25ms (as is appropriate for short-term acoustic features, Tzanetakis & Cook, 2002), with a 50% overlap. To investigate the phase relationship with spectral flux and each of the cardiac and respiration measures, we first calculated the average frequency of heart cycles and respiration cycles, which were 1.01 Hz and 0.30 Hz, respectively. Spectral flux was transformed into the frequency domain over these frequencies using 3-cycle complex Morlet wavelets with centre frequencies of 1.01 Hz and 0.3 Hz. Finally, the angle was extracted with the *angle* function in MatLab. For correlations between heart/respiration rate and spectral flux, spectral flux was resampled to account for slower rate of change of heart/respiration rate. Therefore, to allow meaningful correlations, spectral flux was resampled at 1 Hz to fit more appropriately with heart and respiration rate, then resampled at 20 Hz, i.e., the sampling rate of HR and RR (see below '*Time Domain synchrony*').

The section boundaries in the musical pieces were identified either by a double bar line or end repeat bar line, or by a change/repeat of thematic material (by AC, then confirmed by a music theorist, see Supplementary Table 1).

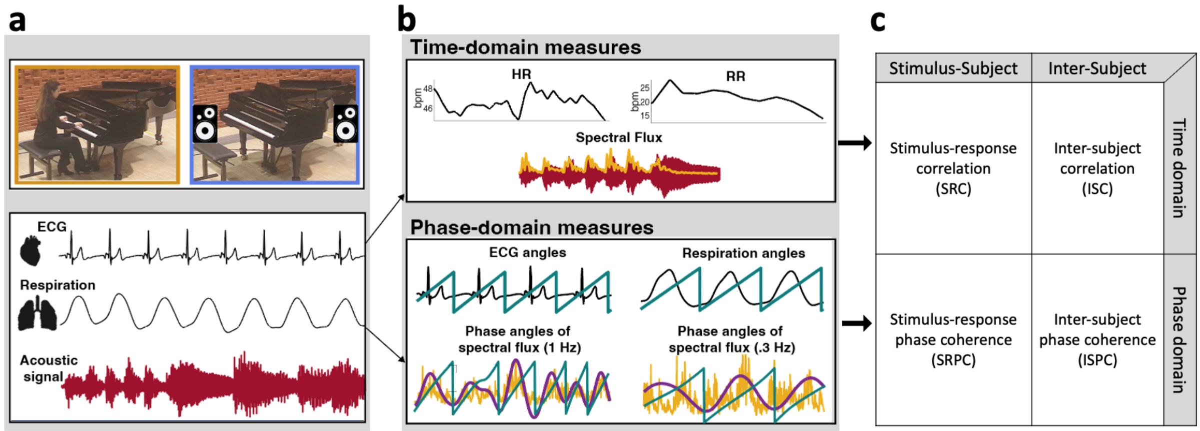


Figure 1. Outline of experiment and analysis pipeline. **A** shows the study design: audiences were presented with music pieces in two conditions audio-only (AO, blue) or audio-visual (AV, orange) condition, where heart (ECG), respiration, and acoustic signal of music (maroon) was continuously recorded. The musical pieces were: Johann Sebastian Bach: Prelude and Fugue in D major (Book Two from the Well-Tempered Clavier, BWV 874), Ludwig van Beethoven: Sonata No. 7, Op. 10, No. 3, second movement (Largo e mesto), and Olivier Messiaen: *Regard de l'Esprit de joie* (No. 10 from *Vingt Regards sur l'enfant Jésus*). **B** shows the time domain measures (upper panel) and the phase-domain measures (lower panel) we extracted from the continuous measures. For the time-domain measures, we extracted heart and respiration rate (HR, RR, from cardiorespiratory signals) and spectral flux (from acoustic signal). For the phase domain measures, we extracted phase angle of ECG and respiration cycles (from cardiorespiratory signals) and angles of spectral flux in the frequency band of heart frequency (1 Hz) and respiration frequency (.3 Hz) (from acoustic signal). In **C,** we show the synchrony measures extracted: stimulus-subject (vertical, left) and inter-subject (vertical, right) in the time (horizontal, top) and phase domains (horizontal, bottom).

Physiological pre-processing.

Heart and respiration signals were pre-processed in MATLAB 2019b. Any missing data (gaps < 53 ms) from the raw signals were first interpolated at the original sampling rate. Data were cut per piece and further preprocessed in MATLAB using Fieldtrip (Oostenveld et al., 2011) and the *biosig* toolbox (http://biosig.sourceforge.net/help/index.html). Respiration data were low pass filtered at 2 Hz, ECG data were band-pass filtered between 0.6 and 20 Hz (Butterworth, 4th order), and both demeaned. Peaks in ECG and respiration signals were extracted using, respectively, *nqrsdetect* function from biosignal and a custom-made script that located when the respiration signal exceeded a peak threshold.

Computationally identified peaks were manually screened to ensure correct identification of QRS and respiration peak locations. Any peaks that were not correctly identified were manually added, while falsely identified peaks were removed, for example, if a T wave in the ECG was accidentally identified as a QRS peak. Data that were too noisy for the identification of clear QRS/respiration peaks were rejected from further analysis (ECG = 14%, respiration = 7%).

Synchrony analyses

Four kinds of continuous synchrony were extracted: stimulus-response correlation (SRC), intersubject correlation (ISC), stimulus-responses phase coherence (SRPC), and intersubject phase coherence (ISPC) for each ECG and respiration, yielding eight kinds of synchrony measures in total.

Time domain synchrony: ISC and SRC

Heart rate (HR) and respiration rate (RR) were calculated by obtaining the differential timing between peaks, i.e., inter-beat-intervals (IBI) for ECG, and inter-breath intervals-(IBrI) for respiration. IB(r)Is were then converted to beats per minute (bpm) and interpolated at the original sampling rate to obtain instantaneous HR and RR, which were down sampled to 20 Hz (Merrill et al., 2021). SRC and ISC were then calculated in a temporal fashion. Although these measures have previously been calculated at 5 second sliding window for EEG (Dauer et al., 2021; Kaneshiro et al., 2019), peripheral responses of heart and respiration are slower moving and take up to 10 seconds for an evoked response (Bradley & Lang, 2000; Sammler et al., 2007; cf. Czepiel et al., 2021). Therefore, we chose a 10 second sliding time window with a one second overlap.

For SRC, each participants' HR and RR signal was correlated with spectral flux. To account for lags in responses, we adjusted data with a stable lag to optimally align the stimulus and the corresponding responses. To find optimal stable lags for each respiration and heart rate, we first calculated each individual's maximum IBI/IBrI (on average, heart: 1000 ms; respiration: 5800 ms), as it might take up to one heartbeat/respiration cycle to respond to the stimulus. SRC at lags up to this maximum IBI/IBrI using *xcorr* in MatLab were calculated and assessed lag at peak correlation (Luck et al., 2010). This was done within the first 10 seconds of the stimulus onset, as this initial onset of the stimulus likely evokes the most reliable response. The average optimal lag for heart and respiration rate were 100 and 300 ms, respectively. These stable lags were applied for all heart and respiration stimulus-response pairings.

For ISC, $p \times t$ matrices were created, where p was the HR/RR signal for each participant over t timepoints. Signals were correlated in a pairwise fashion, i.e., between all possible participant pairs within one concert. Values underwent Fisher-Z transformation and were averaged and then transformed back to r values (inverse Fisher-Z).

Phase domain synchrony: ISPC and SRPC

To transform cardiorespiratory signals to the phase domain, first a continuous sinusoidal wave was fitted to detrended and normalised ECG and respiration signals (based on (Assaneo et al., 2021) by fitting a sine wave to each IBI, using the following equation:

$$A \sin\left(2\pi f_{(k)} t_{(k)} + \theta\right)$$

where A is the mean peak amplitude (i.e., average amplitude of signal at time points of QRS / respiration peak), $f_{(k)}$ is the frequency calculated from IBI (i.e., peak_{k+1} -peak_{k}), converted to Hz ($\frac{1}{IBI(s)}$). Phase (θ) was optimised so the peak of the sine wave corresponded to the ECG/respiration peak. Next, the phase angle was calculated from the real part of the Hilbert envelope of this sine wave using the angle function in MatLab.

SRPC was obtained by calculating the continuous phase coherence between the ECG and respiration phase angles with the spectral flux phase angles corresponding to ECG frequency (1.01 Hz) and respiration frequency (0.30 Hz). Phase coherence between these two signals was calculated based on intertrial phase clustering, with the following formula based on Cohen (2014):

$$ISPC = \left| n^{-1} \sum_{r=1}^{n} e^{ik_{tr}} \right|$$

where n is the number of phase signals, e^{ik} is Eulers formula (i.e., the complex polar representation of phase angle k for signal r at time point t. To account for lag between the two signals, similar to SRC, we applied a stable lag to optimally align phase (van den Bosch der Nederlanden, 2020). To find optimal stable lags, we calculated the maximum cycle length, i.e., maximum IBI/IBrI (on average, heart: 1000 ms; respiration: 5800 ms) for each individual, and calculated phase coherence at lags between 0 seconds until this maximum length at 10 ms steps. Coherence values were then calculated at the optimal lag (i.e., lag with highest coherence value). Similar to SRC, this was done in the first 10 seconds of the stimulus onset, as this initial onset of the stimulus likely evokes the most reliable response. The average optimal lag for heart and respiration phase were 210 and 2540 ms respectively. These lags were then applied for all heart and respiration stimulus-response pairings.

ISPC was assessed in a pairwise fashion between the ECG and respiration phase angles of each participant and the angle of each of the rest of the participants within that concert to obtain a continuous ISPC using the same coherence formula as above.

Global and temporal synchrony

Synchrony measures (SRC, SRPC, ISC, ISPC for each ECG and respiration) were calculated in a global and time-resolved fashion to test the different hypotheses. For the first two hypotheses, we calculated global synchrony values. Continuous synchrony values within each piece were cut into seven, nine, and nine sections, for Bach, Beethoven, and Messiaen, respectively (for details see Supplementary Table 1) and then averaged over these piece sections. For the third hypothesis, we calculated temporal synchrony values by cutting epochs ±15 seconds relative to structural boundary onsets, to capture event-related respiration and heart responses (Bradley et al., 2001; Sammler et al., 2007), as well as any anticipatory effects at musical events (Wassiliwizky et al., 2017).

Significance

To assess synchrony significance (first hypothesis), we tested the assumption that responses time-locked to a stimulus should evoke similar responses across participants (Nastase et al., 2019). Therefore, our 'control' condition was time-'unlocked' data, created by circular shifting ECG and respiration signals. This was done 1000 times and global synchrony measures were calculated as above for each of the shifted data sets. Following Harding et al. (2019), one true data value and one permutated value (the average) per stimulus and participant was obtained. As data were not normally distributed, a Wilcoxon signed rank test assessed the significance across aggregated synchrony values. As we had eight synchrony measures, we corrected the alpha level to .05 / 8 = .006.

AV vs. AO: global

To test the second hypothesis, global synchrony measures (SRC, SRPC, ISC and ISPC) were compared between modalities. Linear mixed models (LMMs) with fixed effect of modality were constructed for each of the synchrony measures as the outcome variables. As the design meant that data was clustered, we added random effects to account for this non-independence. Random intercepts for concert, piece, and participants were added; participants were nested within concerts, while participant and piece were considered crossed effects. We also included a random slope for participants. This random-effect structure represents maximal models (Barr et al., 2013; Page-Gould, 2017). If these maximal models generated convergence and/or singularity fit errors, models were simplified following recommendations (Barr et al., 2013; Barr, 2021). As output from models generating errors should not to be trusted (Barr et al., 2013), we present simplified, error-free models in the results. Maximal model outputs can be nonetheless found in Supplementary Materials. LMMs were run using *lmer* from the *lme4* packages (Bates et al., 2015; Kuznetsova et al., 2017). Significance values and effect sizes, were obtained from the *tab_model* function from *sjPlot* package (Lüdecke, 2023).

AV vs. AO: temporal

To test the third hypothesis, synchrony trajectories at structural boundaries were assessed between modalities across time windows. Synchrony trajectories were split into time windows of 4 seconds. Linear mixed models (LMMs) with fixed effect of modality and time window were constructed for each of the aggregated synchrony trajectories as the outcome variables. As above, we added random intercepts for concert, piece, and participants (participants nested within concerts; participant and piece as crossed effects) and a random slope for participants. Estimated marginal means (using *emmeans* package, Lenth et al., 2018) were used to check pairwise comparisons with Bonferroni adjustments. We expected that synchrony would be highest in the time window centred at the structural boundary. Therefore, we expected this centred time window to have significantly higher synchrony values than other time windows.

3 Results

SRC and ISC significant

Time domain synchrony measures of SRC and ISC were significant compared to circular shifted control for both ECG and respiration with large effect size (see Table 1, Cohen 1988). We therefore only take SRC and ISC into further analysis to assess whether these synchronies were related to modality influences at global and temporal levels.

Table 1. Results from Wilcoxon signed-rank tests between synchrony calculated with original (time-locked) data and synchrony calculated with control (circular shifted) (calculated 1000 times).

	Heart			Respiration		
	Wilcoxon signed rank statistic	**p**	**Effect size**	**Wilcoxon signed rank statistic**	**p**	**Effect size**
SRC	**2**	**< .001**	**0.863**	**34**	**< .001**	**0.677**
ISC	**13**	**< .001**	**0.793**	**16**	**< .001**	**0.782**
SRPC	86	.0119	0.330	112	.278	.225
ISPC	127	.754	0.070	119	.39	0.181

Global Synchrony between modalities

The AV condition evoked higher SRC and ISC for both ECG and respiration. This was significant only for ISC-HR (see Table 2). Confirming the behavioural relevance of synchrony as an engagement index, self-reports in immersion (German: ‘versunken’) (ß = 0.46, 95% CI [0.03 0.90], $p = .038$, $R^2_{(fixed)} = 0.019$) were also significantly higher in AV.

Table 2. Linear mixed models for SRC and ISC.

Linear mixed models comparing synchrony between modalities: SRC

	SRC-HR			**SRC-RR**		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
(Intercept)	0.044	0.020 – 0.069	**<0.001**	0.022	-0.001 – 0.045	0.066
mod [AV]	0.008	-0.006 – 0.022	0.281	0.010	-0.004 – 0.024	0.167
Random Effects						
σ^2	0.01			0.01		
τ_{00}	$0.00_{\text{mac:conc}}$			$0.00_{\text{mac:conc}}$		
	0.00_{piece}			0.00_{piece}		
τ_{11}	$0.00_{\text{mac1.modAO}}$					
	$0.00_{\text{mac2.modAV}}$					
ρ_{01}						
ρ_{01}						
ICC	0.07			0.05		
N	3_{piece}			3_{piece}		
	16_{mac}			16_{mac}		
	2_{conc}			2_{conc}		
Observations	1075			1152		
Marginal R^2 / Conditional R^2	0.001 / 0.076			0.002 / 0.055		

Linear mixed models comparing synchrony between modalities: ISC

	ISC-HR			**ISC-RR**		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
(Intercept)	0.015	0.003 – 0.027	**0.013**	0.034	0.017 – 0.050	**<0.001**
mod [AV]	0.012	0.002 – 0.022	**0.024**	0.004	-0.011 – 0.019	0.590
Random Effects						
σ^2	0.00			0.01		
τ_{00}	0.00_{piece}			$0.00_{\text{mac:conc}}$		
				0.00_{piece}		
				0.00_{conc}		
τ_{11}	$0.00_{\text{mac1.modAO}}$					
	$0.00_{\text{mac2.modAV}}$					
ρ_{01}						
ρ_{01}						
ICC	0.05			0.01		
N	3_{piece}			2_{conc}		
	16_{mac}			3_{piece}		
				16_{mac}		
Observations	1075			1010		
Marginal R^2 / Conditional R^2	0.007 / 0.054			0.000 / 0.009		

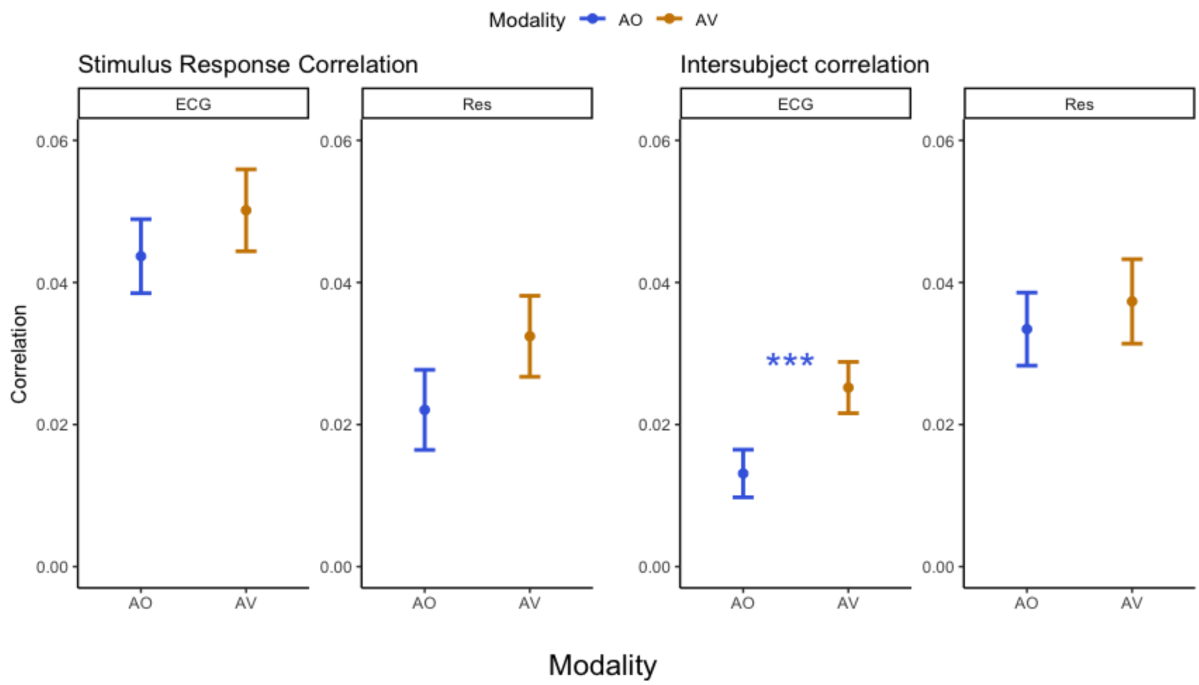


Figure 2. Modality differences for time domain synchrony measures.

Synchrony increases at section boundaries, depending on modality

F-tests for LMM and pairwise comparisons for synchrony measures in time windows at structural boundaries are shown in Table 3 and 4. Time window was a significant predictor for all synchrony measures at structural boundaries, where synchrony increased at and just before the structural boundary for SRC and ISC for both HR and RR (see Figure 3). This increase was confirmed by pairwise comparisons showing that synchrony values were indeed significantly higher in the time window centred at structural boundary (window 0, see Table 3) compared to time windows before (12 seconds before, ISC-HR), and after (4 seconds after, SRC-HR, ISC-HR; 8 seconds after SRC-HR, ISC-HR; and 12 seconds after, SRC-HR, SRC-RR and ISC-HR) in most synchrony measures. Additionally, some synchrony measures (SRC-HR, ISC-RR) were significantly higher in the time window just before at structural boundary (window -4) compared to time window at structural boundary (window 0). When checking what kind of responses were more highly synchronised at the time points (cf. Czepiel et al., 2021), we see that there are synchronised deceleration-acceleration HR and RR responses that significantly decrease and increase over time (see Supplementary Material).

There was also an interaction of modality and time window for SRC-HR and ISC-HR. Figure 3 shows that synchrony was higher just before (window -4) and at structural boundaries (window 0), especially in AV condition. This was confirmed by significant pairwise comparisons for ISC-HR.

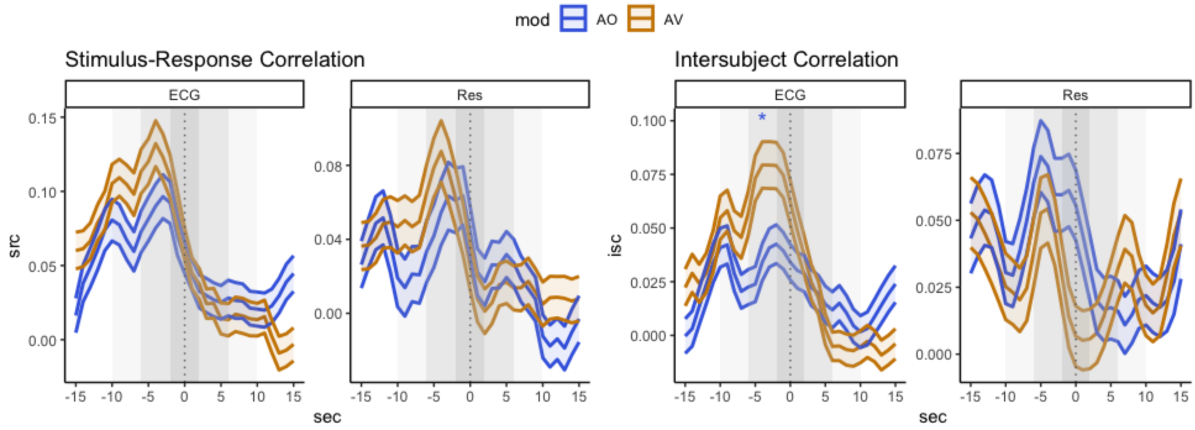


Figure 3. Synchrony values across time centred at structural boundaries. Significance stars in the graph (blue) within a time window represent a significant modality effect within that time window.

Table 3. F-tests for linear mixed effects models for synchrony measures with predictors of time window and modality.

		Heart rate				Respiration rate			
		F	df	p	R (m)	F	df	p	R (m)
	Modality	0.701	(1, 13)	.144		2.510	(1, 934)	.113	
SRC	Window	**18.640**	**(6, 850)**	**>.001**	0.109	**5.953**	**(6,924)**	**>.001**	0.042
	Modality x window	2.300	(6, 850)	.033		1.267	(6,924)	.270	
	Modality	3.524	(1,15)	.079		1.242	(1, 937)	.26	
ISC	Window	**14.560**	**(6, 856)**	**>.001**	0.104	**3.453**	**(6, 925)**	**.002**	0.030
	Modality x window	**4.478**	**(6, 856)**	**>.001**		1.650	(6, 925)	.156	

Table 4. Pairwise comparisons from significant effects in linear mixed models.

			ß	SE	t	p
SRC-HR	Window	Win 0 – Win-12	0.002	0.011	0.191	1.00
		Win 0 – Win-8	-0.023	0.011	-2.066	.235
		Win 0 – Win-4	**-0.038**	**0.011**	**-3.432**	**.004**
		Win 0 – Win 4	**0.040**	**0.011**	**3.556**	**.002**
		Win 0 – Win 8	**0.044**	**0.011**	**3.950**	**< .001**
		Win 0 – Win12	**0.047**	**0.011**	**4.170**	**< .001**
	Modality x window	win-12 AO – win-12 AV	-0.02584	0.0155	-1.666	0.6723
		win-8 AO – win-8 AV	-0.03152	0.0155	-2.033	0.2967
		win-4 AO – win-4 AV	-0.03622	0.0155	-2.336	0.1382
		win0 AO – win0 AV	-0.01033	0.0155	-0.666	1.0000
		Win4 AO – win4 AV	0.00860	0.0155	0.555	1.0000
		Win8 AO – win8 AV	0.00763	0.0155	0.492	1.0000
		win12 AO - win12 AV	0.02739	0.0155	1.766	0.5438
SRC-RR	window	Win 0 – Win -12	-0.009	0.012	-0.728	1.000
		Win 0 – Win -8	0.002	0.012	0.141	1.000
		Win 0 – Win -4	-0.025	0.012	-2.073	.23
		Win 0 – Win 4	0.014	0.012	1.171	1.000
		Win 0 – Win 8	0.024	0.012	2.031	.225
		Win 0 – win12	**0.036**	**0.012**	**3.061**	**.014**
ISC-HR	Window	**win0 – (win-12)**	**0.026**	**0.007**	**3.6545**	**0.003**
		win0 – (win-8)	0.007	0.007	0.895	1.00
		win0 – (win-4)	-0.001	0.007	-0.120	1.00
		win0 – win4	**0.033**	**0.007**	**4.440**	**<.001**
		win0 – win8	**0.046**	**0.007**	**6.17**	**<.001**
		win0 – win12	**0.044**	**0.007**	**5.850**	**<.001**
	Modality x window	win-12 AO – win-12 AV	-0.01533	0.010	-1.299	1.00
		win-8 AO – win-8 AV	-0.017	0.010	-1.476	0.987
		win-4 AO – win-4 AV	**-0.043**	**0.010**	**-3.080**	**0.001**
		win0 AO – win0 AV	-0.029	0.010	-2,545	0.080
		Win4 AO – win4 AV	0.00711	0.010	0.285	1.00
		Win8 AO – win8 AV	0.01241	0.010	1.285	1.00
		win12 AO – win12 AV	0.01716	0.010	1.728	0.595
ISC-RR		win0 – (win-12)	-0.018	0.01	-1.677	1.00
		win0 – (win-8)	-0.001	0.01	-0.150	1.00
		win0 – (win-4)	**-0.029**	**0.01**	**-2.713**	**0.041**
		win0 – win4	0.011	0.01	1.012	1.00
		win0 – win8	0.004	0.01	0.408	1.00
		win0 – win12	0.003	0.01	0.315	1.00

4 Discussion

An important aspect of a music concert is the visual information of seeing musicians perform (Wald-Fuhrmann et al., 2021). The current study investigated whether concert audiences' engagement – indexed here by cardiorespiratory synchrony – would be higher during audio-visual (AV) than in audio-only (AO) music performances. Participants listened to AO and AV performances of Western classical music in a concert setting while cardiorespiratory measures were continuously measured. First, results revealed that time domain cardiorespiratory synchrony measures were significant, i.e., significantly higher compared to control (circular-shifted) data. Second, ISC-HR and self-reports of immersion were significantly higher in the AV condition than in the AO music performances. Third, synchrony differences between modality were investigated on a local level at salient moments in the music, operationalised here as structural boundaries. Synchrony measures significantly increased prior to a structural boundary especially in AV performances, shown significantly in ISC-HR.

Time-locked synchrony between participant responses, and to a stimulus, has been found to reflect engagement to auditory stimuli (Dmochowski et al., 2012; Herrmann & Johnsrude, 2020; Irsik et al., 2022; Kaneshiro et al., 2020; Madsen et al., 2019). So far, these results mostly concern neural measures from studies conducted in laboratory settings. With frameworks on the one hand proposing to extend our understanding of neural mechanisms to peripheral rhythms (Criscuolo et al., 2022; Klimesch, 2018; Parviainen et al., 2022), and on the other to conduct research in more ecological settings (Tervaniemi, 2023; Wald-Fuhrmann et al., 2021), the current study explored cardiorespiratory synchrony in real-world concerts. We calculated four kinds of cardiorespiratory synchrony: stimulus-response and inter-subject synchrony in both time and phase domains (Kaneshiro et al., 2020; Weineck et al., 2022, see Figure 1).

In assuming that non-time locked responses would not be correlated as nothing 'couples' them (Madsen et al., 2019, 2022; Pérez et al., 2021), we first assessed significance of cardiorespiratory synchrony against a time 'unlocked' control data set (i.e., circular shifted data, Madsen et al., 2019). Results reveal that SRC and ISC for both HR and RR were significant. These results indicate that stimulus-related synchrony in neural measures (Dmochowski et al., 2012; Kaneshiro et al., 2020) can also be observed in heart rate (Madsen et al., 2022; Pérez et al., 2021). We additionally provide evidence that synchrony of respiration rate can also occur in music listening. This contrasts Madsen et al. (2022), who did not find ISC-RR was significantly higher than control data. One explanation why the current study did find that ISC-RR is significant is perhaps the difference in stimuli. We used musical stimuli, whereas Madsen et al. (2022) used speech (instructional videos). This suggests that there might be differences in synchrony in autonomic signals, depending on whether the environmental input is music or speech. However, a more convincing interpretation could be that the audiences here were much more engaged with the stimuli used here than in Madsen (2022). In particular, music could be more engaging than instructional videos. More importantly, the setting of a concert hall might have increased the engagement of the listeners compared to the laboratory setting. This therefore highlights that cardiorespiratory synchrony may be context dependent.

SRPC of both heart and respiration cycles was not significant. This was expected for the heart rhythms, as previous work showed that the heartbeat does not align with external beats (Ellis & Thayer, 2010; Koelsch & Jäncke, 2015; Mütze et al., 2020). However, this was unexpected for respiration, as previous studies found that breathing aligns with musical beats (Etzel et al., 2006; Haas et al., 1986). A reason for this discrepancy could relate to the rhythmic structure of the music. These previous studies used music all with relatively regular rhythmic and metric structures (Etzel et al., 2005; Haas, 1986). However, one of the stimuli used in the current study was a contemporary piece with ambiguous harmony. When checking piece differences (see Supplementary Materials), there were no significant synchrony differences between pieces with regular (Bach, Beethoven) and irregular (Messiaen) metric structures. Thus, this explanation of rhythmic structure differences can be ruled out. A more convincing explanation are differences between laboratory and real-world contexts. It seems more intuitive that respiration would align in a concert context as there is likely more immersion in the music compared to a laboratory setting. However, respiration is perhaps more likely to align with beats in a lab as the focus is mainly on music. In a real-world setting, there are several other factors next to the music, such as the visual and social aspect (Wald-Fuhrmann et al., 2021). These additional real-world factors are more representative of a real-world listening situation and might reduce the ability to align breathing to the musical beat. Thus, although neural (Harding et al., 2019; Vanden Bosch der Nederlanden et al., 2020) and respiratory rhythms (Etzel et al., 2005; Haas, 1986) may align to music in a laboratory setting, the current results suggest this does not occur in a music concert setting. This then also highlights the differences between time-domain and phase-domain synchronies. Time domain might be more likely reflecting engagement (see paragraph above), while phase-coherence might be indeed more related to intelligibility to stimulus (Harding et al., 2019; Luo & Poeppel, 2007; Peelle et al., 2013; Vanden Bosch der Nederlanden et al., 2020).

There were also no significant effects for ISPC of heart and respiratory cycles. Perhaps this is unsurprising as phase synchrony between participants has previously been found only in situations of social interaction (Hartmann et al., 2019; Richardson et al., 2012; Toiviainen & Carlson, 2022) to achieve a specific goal, such as duets (Gugnowska et al., 2022; Lindenberger et al., 2009) and ensemble performances (Lindenberger et al., 2009; Vickhoff et al., 2013). In a typical Western classical concert, the etiquette is to avoid any communication or interaction between audience members, in order to enhance the focus on music (Wald-Fuhrmann et al., 2021). Indeed, from checking the videos, the audience members respected this etiquette and were not directly interacting with each other. Therefore, we corroborate the idea that phase synchrony of heart and respiratory cycles - which is typically seen in interactive settings - may not occur in non-interactive settings.

Next, significant synchrony measures were compared between modality conditions. Both ISC and SRC synchrony measures were greater in AV than AO condition, but significantly so in ISC-HR only. As self-reports of immersion were likewise higher in the AV condition, this suggests that ISC-HR may reflect higher engagement with the stimuli. This is in line with Ki et al. (2016) who found engagement effects, as reflected in neural ISC, were stronger in AV compared to AO stimuli. This additionally supports studies showing that ISC of EEG is higher during attentive than distracted music listening (Madsen et al., 2019). In extending these previous results found in neural measures, the current results contribute to growing evidence that heart rate synchrony can reflect engagement not only to narratives (Pérez et al., 2019), but also to music. This heart rate synchrony across audience members might reflect heightened immersion in music, as there is more similar processing when listeners attended more to the music. Overall, the increase of ISC-HR and immersion in AV compared to AO suggest that the audience had higher engagement when they could see the musician in the performance.

Finally, synchrony changes between modalities at salient moments in the music were assessed. Salient music moments were defined here as structural boundaries, as they typically are locations when several music features change simultaneously (Clarke & Krumhansl, 1990; Deliege, 1987; Krumhansl, 1996; Phillips et al., 2020; Popescu et al., 2021). Here, both ISC and SRC increased at structural boundaries, supporting and extending previous work on ISC (Czepiel et al., 2021; Dauer et al., 2021) to SRC. As attention with music typically fluctuates across time (Fink et al., 2018), such increases of cardiorespiratory synchrony at structural boundaries found here suggests increased engagement with music at these time points. We corroborate this by showing that the synchronised responses were deceleration-acceleration HR and RR patterns, reminiscent of arousal (Boiten et al., 1994) and orienting responses (Bradley & Lang, 2000; Park et al., 2014; Proverbio et al., 2015; Stekelenburg & Boxtel, 2001). Auditory cues in the music that indicate a section is ending, such as cadential ending (i.e., harmonic closure in Bach and Beethoven) and a slight slowing down or change in tempo and/or material (see Supplementary Table 1), likely causing participants bring back their attention to the music. These physiological patterns at salient changes in musical features suggest a bottom-up capturing of engagement with the music.

In terms of modality differences at structural boundaries, Figure 3 shows that most synchrony measures increase at structural boundaries more so in the AV conditions, significant only for ISC-HR. The ISC-HR increase in the AV condition suggests that visual information further increases engagement to the structural boundary than just audio information alone. Previous performance studies have found that certain gestures cue and anticipate structurally important locations, such as leaning forward towards a cadence (harmonic closure) (Davidson, 2012), increased movement amplitude (Thompson & Luck, 2012) or finishing 'flourish' gestures (Wanderley et al., 2005), as well as deep breaths and swooping motions prior to onset of new phrases (Vines et al., 2004). Indeed, videos show that the pianist tended to lean forward at structural boundaries. Therefore, a combination of audio and visual aspects likely increase engagement with music on a local level, as shown most robustly with ISC-HR.

5 Limitations

Although we interpret greater synchrony in AV (compared to AO) as engagement, it is unclear whether this synchrony increase might also be related simply to increased information processing. However, as we did not have a condition which was visual only, future work would be required to disentangle the synchrony differences between additional information processing and actual engagement.

Another limitation of the study is the potential confounding variable that AV was presented in a live version, while the AO condition was presented as a recording. This was chosen in order to enhance ecological validity: people who listen to music in an AO version most likely listen to music as a recording, while watching an AV version is more likely to be live (Sloboda et al., 2012). However, we appreciate that today people also watch recorded music in the audio-visual modality, using tools such as YouTube, Digital Concert Hall of the Berlin Philharmonic, and MetOnDemand. Nonetheless, recent work has shown that there has been little to no difference between recorded and live version of music in terms of pleasantness ratings (Belfi et al., 2021). Despite this, future work would have to be mindful of this and control for liveness.

6 Conclusion

Using cardiorespiratory synchrony to index engagement (Dmochowski et al., 2012; Kaneshiro et al., 2020; Madsen et al., 2019), the current study explored engagement between audio-visual and audio-only music performances in a Western classical concert setting. Extending studies investigating neural synchrony (Dmochowski et al., 2012; Kaneshiro et al., 2020) this study contributes to growing literature on peripheral physiological synchrony (Madsen & Parra, 2022; Pérez et al., 2021). Time domain cardiorespiratory synchrony was greater when audiences saw the musician, both at a global and time-resolved level. Such synchrony increases suggest that AV musical performance evoked higher engagement in listeners. This higher engagement may be a reason why people attend such events. Simple everyday music listening, which is commonly background listening of audio-only recordings, may be less engaging (Sloboda & O'Neill, 2001). However, people can more actively engage with music in a live concert setting, which might subsequently enhance engagement and enjoyment (Koelsch et al., 2019). ISC-HR was the most robust synchrony measure, informing future studies about which measures might be most informative to study musical engagement in ecological settings. We found no significance of phase synchrony. Phase synchrony has previously been found when participants can move and interact with each other (Gugnowska et al., 2022; Hartmann et al., 2019; Toiviainen & Carlson, 2022) in music contexts. The fact we do not replicate such results here might be due to the fact that typical Western classical concert etiquette requires audience members to limit movements and interactions. Thus, questions arise whether phase coherence might become more relevant in contexts where listeners can move and interact, for example in pop and rock concerts. Overall, more studies from wider-ranging genres are needed to further understand synchrony dynamics in real-world music contexts.

Acknowledgements

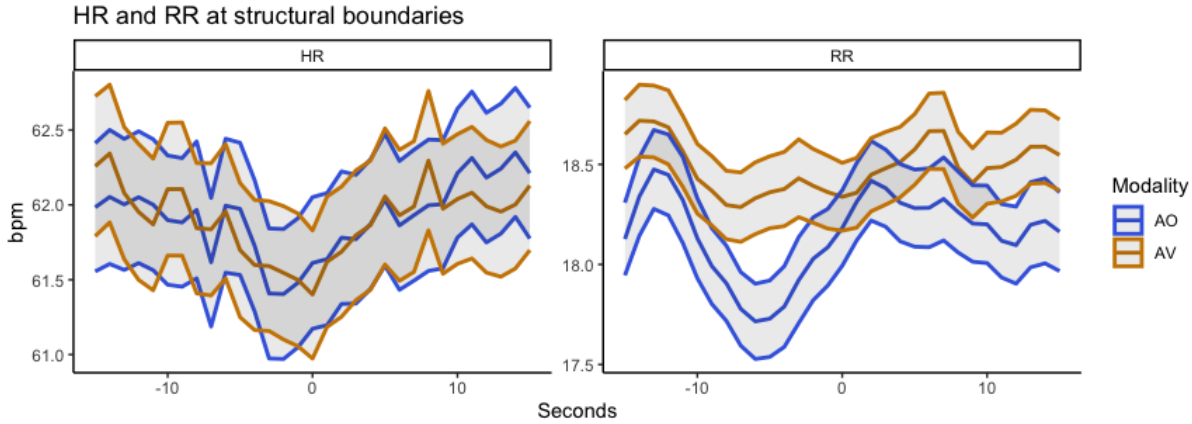
The authors would like to thank Lea Fink, our music theorist who double checked the musical section boundaries. Thanks also go to Klaus Frieler and Örjan De Manzano for discussions about statistics. Many thanks also to the ArtLab team, who assisted in concert preparations.

7 Supplementary Materials

Supplementary Table 1: Outline of the section boundaries. For Bach's Prelude and Fugue, the sections were clear with double bar lines, repeat bar lines, or a change in thematic material, tempo and/or harmony. For Beethoven's Sonata, section changes occurred mainly with changing material and/or double bar line. For Messiaen, section changes occurred with changes of tempo (indicated by the composer) and/or double bar lines. Supplementary Table 1 presents a detailed description of when each section occurs.

Piece	Section	Bar	Metre	Tempo		Description
				M	SD	
Bach	1	1	12/8	88	5	First section of prelude
	2	17	12/8	86	4	Repeat of first section of prelude
	3	33	12/8	84	4	Second section
	4	57	12/8	86	3	Introductory material in second section
	5	73	12/8	84	4	Repeat of second section
	6	97	12/8	82	11	Repeat of introductory material in second section
	7	113	2/2	88	9	Fugue
Beethoven	1	1	6/8	48	3	Quiet, relatively clear quaver pulse
	2	9	6/8	54	5	Bass has semi quavers Alberti bass
	3	17	6/8	50	7	Dynamics changing loud, forte, reinforced chords
	4	26	6/8	51	5	Quaver pulse in left hand, nice quiet melody in right hand, which gets louder over time
	5	30	6/8	58	9	Quaver pulse, major low pitch right hand, followed by minor, then but high-pitched demisemiquavers in right hand
	6	44	6/8	58	5	Introductory material again
	7	56	6/8	53	6	Repeat of section 43 material
	8	65	6/8	50	8	Quaver pulse LH, RH semidemi triplets, which speed up to hemidemisemiquavers, starts quieter and lower in register, increase in pitch and dynamic.
	9	76	6/8	58	10	Closing section. Quaver pulse mostly. Semitone intervals
Messiaen	1	1		86	8	Pres vif (quaver = 160). Pulsating semiquavers, occasional 'blasts' of acciaccatura quaver.
	2	2		60	11	Modéré (quaver = 138). Expressing
	3	3		71	15	Un peu plus vif (quaver = 160). Quaver trills in H, RH semiquaver triplet
		4		71	17	Bien modéré (mais de plus en plus véhément crotchet tied to semiquaver = 58. Repetitive 5 semi quaver, (2+2+1) falling motif, occasionally interrupted by accented slower semiquaver chords

		5		42	10	Très modéré, Tempo rubato (quaver = 104). Accented 4 note chords in an arching kind of manner
		6		70	10	Modéré (quaver= 138). Quaver, preceded by 5 semiquaver, and burst scalic upward patterns, followed by semiquaver pattern
		7		41	20	Similar to section 5
		8		81	14	Similar to section 1
		9		37	29	Starts similar to Section 5 and 7, followed by trills in LH, hemidemisemiquaver pulsations in RH.



Supplementary Figure 1. Trajectories of actual HR and RR data at section boundaries. As previously done in Czepiel et al. (2021), we wanted to assess what kind of responses became synchronised. Therefore, the actual HR and RR were investigated. HR and RR were split into time windows of four seconds. Two LMMs were constructed, each for aggregated HR and RR values as the dependent variable. Both LMMs had predictors of time window and modality and random effects if participant id, piece, and concert. Time was a significant predictor for both HR and RR. As seen in Supplementary Figure 1, there is a HR deceleration, where HR near 0 (i.e., at the structural boundary) is lower than around 10 seconds before structural boundary (nearing significance, $p = .055$). The HR deceleration was followed by a significant HR acceleration between time 0 and around 15 seconds after structural boundary ($p = .021$). This is a typical HR deceleration-acceleration pattern that has been shown to anticipate and perceive events (Bradley et al., 2001; Park et al., 2014). RR significantly decreases prior to section boundary ($p < .023$) and then increases but misses significance (.239). Similar to HR, a RR decrease reflects attentional processes, while the RR increase is a typical response associated with physiological arousal (Bradley et al., 2001; Boiten et al., 1994).

Supplementary Table 2. Testing hypothesis 1 between pieces

		Piece	Wilcoxon signed rank test	Significance
ISC	ECG	**Bach**	**61**	**.035**
		Beet	**12**	**< .001**
		Mess	**22**	**< .001**
	Res	**Bach**	**71**	**.042**
		Beet	**6**	**< .001**
		Mess	83	.056
ISPC	ECG	Bach	148	.495
		Beet	127	1
		Mess	87	.21
	Res	Bach	102	.274
		Beet	147	.944
		Mess	154	.932
SRC	ECG	**Bach**	**19**	**< .001**
		Beet	**2**	**< .001**
		Mess	**17**	**< .001**
	Res	**Bach**	**84**	**.105**
		Beet	**31**	**< .001**
		Mess	**54**	**.008**
SRPC	ECG	Bach	104	.483
		Beet	109	.581
		Mess	98	.371
	Res	Bach	111	.427
		Beet	113	.303
		Mess	144	.864

Supplementary Table 3. Linear mixed model with maximal random-effect structures

Linear mixed models (maximal models) comparing synchrony between modalities

	SRC-HR			SRC-RR			ISC-HR			ISC-RR		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
(Intercept)	0.044	0.020 – 0.069	**<0.001**	0.022	-0.001 – 0.045	0.059	0.015	0.003 – 0.027	**0.013**	0.034	0.016 – 0.052	**<0.001**
mod [AV]	0.008	-0.006 – 0.022	0.280	0.010	-0.006 – 0.026	0.206	0.012	0.002 – 0.022	**0.024**	0.004	-0.013 – 0.020	0.668
Random Effects												
σ^2	0.01			0.01			0.00			0.01		
τ_{00}	$0.00_{\text{mac:conc}}$			$0.00_{\text{mac:conc}}$			$0.00_{\text{mac:conc}}$			$0.00_{\text{mac:conc}}$		
	0.00_{piece}			0.00_{piece}			0.00_{piece}			0.00_{piece}		
	0.00_{conc}			0.00_{conc}			0.00_{conc}			0.00_{conc}		
τ_{11}	$0.00_{\text{mac1.modAO}}$			$0.00_{\text{mac1.modAO}}$			$0.00_{\text{mac1.modAO}}$			$0.00_{\text{mac1.modAO}}$		
	$0.00_{\text{mac2.modAV}}$			$0.00_{\text{mac2.modAV}}$			$0.00_{\text{mac2.modAV}}$			$0.00_{\text{mac2.modAV}}$		
ρ_{01}												
ρ_{01}												
ICC	0.07											
N	2_{conc}			2_{conc}			2_{conc}			2_{conc}		
	3_{piece}			3_{piece}			3_{piece}			3_{piece}		
	16_{mac}			16_{mac}			16_{mac}			16_{mac}		
Observations	1075			1152			1075			1010		
Marginal R^2 / Conditional R^2	0.001 / 0.076			0.002 / NA			0.008 / NA			0.000 / NA		

Chapter 6

Summary and General Discussion

The most pleasurable music listening experiences occur in live concert settings (Gabrielsson & Wik, 2003; Lamont, 2011). One way to capture a listening experience in these contexts with minimal disruption is measuring peripheral responses (Ardizzi et al., 2020; Egermann et al., 2013). However, to date there is limited empirical evidence on how humans experience music in such naturalistic contexts (Brattico, 2021; Tervaniemi, 2023; Wald-Fuhrmann et al., 2021). Therefore, the overall aim of this thesis was to assess a more holistic understanding of a real-world music listening experience in audiences attending Western classical concerts, captured by aesthetic and peripheral responses. In the following, I first summarise the main findings of the empirical studies presented in Chapters 2-5. This summary is followed by a discussion on how the study results inform and extend the (neuro-)cognitive music science field's understanding of aesthetic and peripheral physiological responses of naturalistic music listening. Finally, some limitations and open questions are presented with suggestions for future research.

1 Summary

Empirical models of musical aesthetic experiences generally converge on three components: (1) inputs of music, context, and individual person (e.g., music preference), (2) processing mechanisms, and (3) an aesthetic response output (Anglada-Tort & Skov, 2020; Brattico et al., 2013; Hargreaves & North, 2010; Wald-Fuhrmann et al., 2021). Broadly speaking, research has mainly focused on the second component, with recent studies focusing on naturalistic music processing (Alluri et al., 2012; Burunat et al., 2014; Cheung et al., 2019; Di Liberto et al., 2020; Kern et al., 2022; Omigie et al., 2021). However, less is known about the 'input' components, such as the external environment (Tervaniemi, 2023). Therefore, we examined 'inputs' within a concert context, focussing on two broad aspects. As the concert yields an optimal situation to present full-length, naturalistic music, we first explored naturalistic music itself: the global/local structural and acoustic features of music (Chapters 2-3). Second, next to the auditory experience, the visual aspect of seeing the musician in a musical performance was explored (Chapters 4-5). An additional focus was the 'output', evaluating whether and how physiological responses (Chapter 2 and 4) and synchrony of physiological responses (Chapter 3 and 5) reflect aesthetic responses to music, such as liking (Chapter 2 and 4), emotion (Chapters 2-3), and engagement (Chapters 2 and 5).

Chapter 2 presented an initial glimpse into aesthetic experience trajectories over an entire concert program. So far only two studies have empirically and quantitatively assessed the trajectory over the entirety of an evening performance - within the context of an opera (Balteș & Miu, 2014; Scherer et al., 2019a). Chapter 2 extends their findings to a music concert of instrumental pieces. Across three concerts, audiences were presented with a professionally curated concert program, consisting of a Classical period piece (by Ludwig van Beethoven, 1770-1827), a contemporary classical piece (by Brett Dean, b. 1961), an interval, and finally a Romantic period piece (by Johannes Brahms, 1833-97), all for string quintet. The first and last compositions were common practice period (CCP) pieces, with conventional musical structures, tonal harmony, and regular metres. The middle piece was a contemporary classical piece with non-tonal harmony and irregular meter. The audience reported their experiences on liking, absorption, and emotion response scales. Peripheral responses of skin conductance, heart and respiration rate, and facial muscle activity of individual concert attendees were also collected. First, despite highly negative ratings for the middle contemporary piece, liking ratings for the concert in its entirety was overall high (on average 4.1 out of 5). These high liking ratings suggest that the high positive ratings of the first and last piece may offset the negative ratings for the middle contemporary piece. Second, people's emotional responses were related to the overall structure of the multi-movement works: outer movements yielded more positive responses compared to inner, calmer sections that followed the traditional composition conventions of their time. This is in line with the idea that listeners' emotional trajectories across long-term musical works might reflect tension-resolution across music (Balteș & Miu, 2014). However, there was no manipulation of the order of pieces and their internal structure. Therefore, these first two findings regarding concert programming and emotions responses reflecting musical form can only be cautiously interpreted. Future studies with varied piece and movement order would be required to corroborate initial interpretations offered here. Third, specific physiological responses reflected the overall concert listening experience: heart rate (HR) dropped over time, reflecting an overall calming effect of a concert, but rose at the climactic ending of the final piece. Physiological arousal responses (increased SC and RR) were related to increased ratings of feeling energetic, agitated, and feeling joyfulness, while EMG zygomaticus major (smiling muscle) decreased with liking. These results objectively show that certain experience can be reflected in physiological responses in real-world music listening settings. The findings point to potential components that can be further explored to extend some of the exploratory results found here.

Chapter 3 explored responses to more specific and localised musical features using the same string quintet concert data set used in Chapter 2. Previous studies that assessed responses to musical events have done so in a hypothesis-driven approach, i.e., manipulating or locating arousing musical features/self-reported chills and assessing the corresponding physiological responses. As the stimuli used here were non-manipulated naturalistic music, Chapter 3 presented one of the first studies to explore physiological responses to music in a data-driven approach. Moments of systematic physiological responses were first identified and then musical features at these moments in time were evaluated. Systematic physiological responses were operationalised as moments of high synchrony calculated with inter-subject correlation (ISC), based on the assumption that commonalities between time-locked signals enhance signal-to-noise ratio in stimulus processing (Hasson et al., 2004; Nastase et al., 2019). Musical features were extracted by computational acoustic-based and theoretical score-based methods. Moments of high physiological synchrony revealed higher skin conductance and faster respiration, i.e., indices of activation of the sympathetic nervous system. These synchronised arousal responses were linked to music with faster tempo, suggesting that faster music had evoked higher emotional arousal. The link between faster music evoking higher arousal was confirmed by significant correlations between faster tempi and self-reported high arousal emotions. Additionally, high synchrony moments across three concert audiences occurred at structurally important locations in the music, namely transitions between musical sections, sudden changes in the tempo or key, and when thematic material was repeated, but in an unexpected and altered presentation. In summary, synchrony methods offer a way to assess interindividual music listening experiences in non-experimental, naturalistic paradigms, and here we show that specific musical events may evoke common responses across audience members.

A crucial part of a concert performance is that it engages the listener not only in an auditory experience, but also a visual one. **Chapter 4** explored differences when participants experience audio-only (AO) and audio-visual (AV) musical performances. It is well known that visual information of performer movement enhances the emotional intention (Dahl & Friberg, 2007), expressivity (Davidson, 1993, p. 199; Wanderley et al., 2005), and appreciation of music (for a review, see Platz & Kopiez, 2012). However, knowledge of whether such results extend to a concert setting (Coutinho & Scherer, 2017) is limited. Additionally, there are opposing results as to whether additional sensory information increases (Chapados & Levitin, 2008) or decreases (Vuoskoski et al., 2016) physiological responses during music listening. To examine this further, audiences were presented with audio-visual (AV) and audio-only (AO) versions of piano performances (same music in both conditions). Results presented in Chapter 4 revealed that audio-visual information enhanced the aesthetic experience of the music. For peripheral physiological responses, two main patterns were observed. First, heart rhythm related to activation of the sympathetic nervous system ("fight-or-flight" system) increased during AO performances. The most likely explanation is that the lack of visual information made sound onsets more surprising (Jessen & Kotz, 2011). Second, responses related to motor activity (respiration rate, i.e., respiratory muscles, and facial muscle activity) were higher in the AV performance, suggesting that seeing a musician's movements may have evoked motor mimicry in the audience. However, certain inconsistencies in the statistical models for respiration and muscle activity, further research, with larger sample sizes, is required to confirm results found here. Third, we found the smiling muscle was positively related to aesthetic experience, showing support for the embodied aesthetic theory (Cross, 2011; Freedberg & Gallese, 2007). Overall, results demonstrate that physiological responses may reflect multiple responses, showcasing a more holistic, realistic experience of a music performance.

While **Chapter 3** and **Chapter 4** explored musical features and modality in isolation, **Chapter 5** presented a study exploring these two variables together, that is, assessing responses to between AO/AV modalities across pieces and at musical structural boundaries. Synchrony measures of physiological responses were calculated as a way to assess engagement with music over time (Dauer et al., 2021; Kaneshiro et al., 2020). Building on work from **Chapter 4**, stimulus-response synchrony and inter-subject synchrony were assessed both in the time and phase domain (Kaneshiro et al., 2020; Weineck et al., 2022). In the time domain, HR/RR responses were correlated either with the spectral flux of the acoustic signal (stimulus-response correlation, SRC), or across participants (inter-subject correlation, ISC). In the phase domain, coherence between heart/respiration phase angles and phase angle of spectral flux at heart (1 Hz)/respiration (0.3) Hz frequencies (stimulus-response phase coherence, SRPC) were calculated as well as the coherence between heart/respiration phase angles across participants (inter-subject phase coherence, ISPC). Only time domain measures were significantly higher compared to control (circular-shifted) data. Musical engagement – as indexed by cardiorespiratory synchrony and self-reported immersion in music – was greater in the AV condition, both at a global and time-resolved level. ISC of HR was the most robust synchrony measure, supporting the idea that engagement is not reflected in neural synchrony (Kaneshiro et al., 2020; Ki et al., 2016; Madsen et al., 2019), but also in cardiac synchrony (Pérez et al., 2021). Despite the significance of time domain measures during music listening, we found no significant effect of phase synchrony. This might relate to the interaction between participants. Typical Western classical concert etiquette requires audience members to limit movements and interactions. Contrastingly, phase synchrony has previously been found when participants move to and interact with each other when playing and listening to music (Gugnowska et al., 2022; Hartmann et al., 2019). Therefore, questions arise whether the relevance phase coherence is context dependent. Perhaps phase is less relevant in context with less interaction, such as a Western classical concert, but might be more relevant in contexts where listeners can move and interact, for example in pop and rock concerts. Thus, to gain more insight in synchrony dynamics in real-world music listening, future studies from more wide-ranging genres are required.

2 General discussion

Real-world music listening rarely happens with short clips of music in an isolated small room. Thus, moving beyond the laboratory to more naturalistic contexts is particularly important to better understand more realistic music listening experiences (Brattico et al., 2013; Juslin, 2013; Tervaniemi, 2023; Thompson et al., 2023; Wald-Fuhrmann et al., 2021). Concerts are one optimal context to study a listening experience. Not only is this because much CPP music was purposefully composed for such a public function and context, but also because concerts offer greater opportunities for a special aesthetic experience (Sloboda, 2010) where listeners are likely more focussed on and enjoying music[10]. With these advantages present in a concert setting, it was possible to investigate ecologically valid music listening experiences. Assessing peripheral responses alongside self-reports allowed to objectively and uninterruptedly measure music-evoked emotions, engagement, and liking. In particular, this thesis aimed to answer open questions regarding aesthetic experience of full-length pieces and the interplay of several musical and visual features. This thesis also explored physiological responses to gain insights into potential mechanisms into such responses.

2.1 Naturalistic music listening as a dynamic experience

Previous studies using naturalistic music have focused on pop songs or clips from longer pieces that are restricted to around 4-5 minutes (Dauer et al., 2021; Kaneshiro et al., 2020; Omigie et al., 2021) or single movements from multi-movement works of around 9 minutes (Grewe et al., 2009; Kaneshiro et al., 2021). Albeit a challenging undertaking, this thesis presents an exploration into more realistic music listening across a whole concert program, across 30-minute-long pieces, as well as at local music structural events.

At a concert level, dynamic changes in emotion ratings were interpreted to reflect the order of three, very different, musical compositions (Chapter 2). Piece order is a significant consideration for performers and concert organisers (Blackburn, 2016; Gotham, 2014). Previous qualitative interview studies exploring concert experiences suggest that the order of music pieces plays a crucial role in emotional and aesthetic audience responses (Adams et al., 2023; Pitts & Spencer, 2008). The only empirical research coming close to a concert program explored the repetition of a single piece rather than the order of musical pieces (Halpern et al., 2017). Otherwise, the order of music stimuli in most

[10] Note. Conducting this in a Western classic concert, one also attracts not typical study participants, but participants who enjoy Western classic music (further discussed in Section 3.2).

experimental laboratory studies is randomised to avoid confounding effects. This thesis provides also an initial exploration of concert programming. Findings from Chapter 2 report high positive ratings for the first and last pieces, but high negative ratings for the contemporary piece that was placed in the middle of the concert program. Despite these high negative ratings, when asked at the end of the concert how much participants liked the concert, liking ratings were nonetheless high. Such high liking ratings highlight typical concert programming practice where the more challenging musical piece would be 'sandwiched' between the music that is more familiar and likeable (Blackburn, 2016; Gotham, 2014). Using this ordering, the audience is not 'thrown into the deep end' by listening to the more challenging piece first, nor do they leave the concert with this piece fresh in their memory. Although such an observation was possible within a concert, this dataset did not include an experimental manipulation, i.e., pieces did not change in their order of presentation. If this conclusion were true, we would have expected a drop in the liking ratings of the overall concert if piece order was different. For example, audiences may have given low overall concert ratings if the program ended with the contemporary musical piece. This idea stems from the notion that we tend to remember and evaluate events based on the end (i.e., most recent) part of an event (Fredrickson, 2000; Schaefer et al., 2014). Consequently, an interesting direction for future concert research would be to manipulate the order of musical pieces to confirm the current conclusions.

At an individual piece level, audiences were sensitive to musical structure on a global (Chapter 2) and local (Chapters 3 and 5) level. In terms of global structure, this was related to fluctuation of emotional responses (Chapter 2). Early studies have shown how music can evoke one clear emotion (Khalfa et al., 2008), even when rating emotion dynamically (Krumhansl, 1997). However, more real-world listening often takes place on a much larger scale. Studies attempting to understand such longer-term listening have found that emotions fluctuate across opera performances (up to three hours long, Balteș & Miu, 2014; Scherer et al., 2019). However, a key issue with these opera studies is that an emotional response might not come purely from the music, but from the dramatic narrative present in the sung lyrics and surtitles[11]. Thus, to understand if an emotional change is evoked by the music itself, rather than a narrative, it is also crucial to show this in purely instrumental music i.e., music without words. Here, the CPP pieces followed a typical tension-resolution trajectory across four movements, where harmony develops across the movements and resolves at the end. Inner movements provided a contrasting,

[11] Similar to subtitles for the television, surtitles are translation(s) of the libretto (text) of the opera that are projected above the stage during the performance (Pines, 2002).

calmer emotion compared to more lively outer movements. The current results show that audience emotions changed across the music pieces following the trajectories of such a musical structure. These findings are in line with and extend the results from the opera studies to instrumental music. From this, we can conclude that music has the power to evoke dynamically changing emotions that reflect musical tension-resolution patterns presented across musical pieces.

At a local level, this thesis provided evidence that physiological arousal responses increased at structural boundaries in music. In an exploratory data-driven way, Chapter 3 reports that synchronised physiological arousal was related to structurally important moments of the music. This was replicated by a hypothesis-driven method: Chapter 5 reported that physiological synchrony increased at structural boundaries. Indeed, it had been proposed by the Gestalt principles (Bregman 1990; Deutsch, 1999) and the Generative Theory of Tonal Music (GTTM, Lerdahl & Jackendoff, 1983) that in continuous listening, we 'segment' auditory events. The GTTM proposed that segments are perceived when salient changes[12] occur (Lerdahl & Jackendoff, 1977). When tested empirically, such salient changes were indicated by participants as musical segments in both familiar (Clarke & Krumhansl, 1990; Deliege, 1987; Krumhansl, 1996; Tillmann et al., 1998; Tillmann & Bigand, 2004) and unfamiliar music (Clarke & Krumhansl, 1990; Phillips et al., 2020). However, such conscious behavioural ratings might enhance people's awareness of musical structure and might not be a real representation of how we listen to music. By using peripheral physiology, it was possible to assess realistic music listening, which does not engage the listener in an additional task. The peripheral physiological results (Chapters 3 and 5) extend previous research, suggesting that listeners might perceive musical structure when simply enjoying music.

The current results provide novel evidence showing that emotions and engagement fluctuate over an entire professionally curated program (~90 minutes, Chapter 2), over multi-movement works (~40 minutes, Chapter 2), and at local structural changes in the music (~10 seconds, Chapters 3 and 5). Such engagement and emotional fluctuations might help explain why people enjoy going to concerts. Experiencing to a limited range of musical emotions might be dull. Feeling a variety of contrasting emotions likely enhances interest and engagement. Despite listeners liking contemporary pieces less (Chapter 2), people might be nonetheless willing and interested to try new and challenging works (Silvia

[12] The GTTM suggests a set of hierarchical rules that govern human segmentation, where the perception of group boundaries is characterised by rests, changes in register, or harmonic cadences. When one of these occurs, this might be perceived as a salient change.

et al., 2008), especially if sandwiched between pieces by more well-known and liked composers. Further, participants were sensitive to musical structures. Perception of structure might lead to the segmentation of music, which facilitates the processing of such a complex sound stream of naturalistic music (Lerdahl & Jackendoff, 1983). This increased processing fluency might, in turn, enhance an aesthetic experience (Huron, 2012; Reber et al., 2004). Although full-length naturalistic music has allowed for rich insights into dynamic music listening, due to their design, they remain exploratory. Therefore, it is of interest to complement these exploratory findings with research using appropriate controls, such as comparing music with clear vs ambiguous structural transitions, or comparing music across different levels of (emotional) variation (e.g., a music piece with a typical sonata structure compared with repetitive, Minimalistic music piece). Such controls would allow greater understanding into potential driving forces of emotion and engagement dynamic during music listening.

2.2 Naturalistic music listening: holistic feature representation

An advantage of naturalistic music affords the opportunity to examine how several acoustic features are processed in parallel (Alluri et al., 2012; Burunat et al., 2016), which might collectively contribute to aesthetic responses, such as musical beauty (Omigie et al., 2021). In extending this laboratory work to ecological settings, we also extracted several features that characterised the music and assessed their relationship with synchronised physiological arousal responses (Chapter 3). Previous studies have either used score-based (Sloboda, 1991) symbolic-based (Cheung et al., 2019), or acoustic-based musical characterisation analyses (Alluri et al., 2012; Burunat et al., 2016; Omigie et al., 2021). In attempting a more holistic representation of the music, this thesis presented a first attempt of using both score-based and acoustic-based methods to characterise the music.

For the acoustic-based analysis, features of loudness, timbre, harmonic expectancy, and tempo were computationally extracted. Tempo was the only reliable significant predictor of increased emotional arousal during a chamber music concert experience. These findings support Gomez & Danuser (2007), who report that the rhythmic aspect of tempo and articulation were the main drivers of physiological responses to music. Musically speaking, this is in line with the idea that the structure of multi-movement pieces is broadly based on tempo. Multi-movement pieces generally have fast outer movements with slower internal movements (e.g., fast-slow-fast of a concerto, or fast-slow-minuet-fast of a symphony or quartet, Larue et al., 2001). This might mean that for long-term music listening, physiological emotion is primarily influenced by the key musical element that the global musical structure is based on.

However, it was surprising to find no other acoustic features as significant predictors of physiological emotion. In particular, the current findings are not in line with previous studies that showed physiological arousal to be linked to quantitatively extracted harmonic expectancy. Early works showed that musical pleasure and emotion may arise from musical consonance and dissonance (Blood et al., 1999; Helmholtz & Ellis, 1875), which are used to provide tension and resolution in music (Huron, 2006). Indeed, the expectation and surprise of such harmonic, melodic, and rhythmic structures have been shown as predictors of musically induced emotions and pleasure (Bianco et al., 2019; Cheung et al., 2019; Koelsch et al., 2019). As pleasure has also been related to arousal (Berlyne, 1971), it can be hypothesised that physiological arousal – such as increased skin conductance (SC) – might be related to harmonic and melodic expectations. Indeed, increased SC was linked to surprising harmony and melody in both laboratory studies (Koelsch et al., 2008; Omigie et al., 2021; Steinbeis et al., 2006) as well as in one concert study (Egermann et al., 2013). Chapter 3 did not directly replicate these previous results. There are three potential reasons for this discrepancy. First, moments of physiological arousal were extracted with ISC, a method that highlights commonalities across responses. Although harmonic expectation is implicitly learnt through exposure to music (Koelsch et al., 2000), there may be differences in how individual listeners generate harmonic expectation (Bianco et al., 2018; Loui & Wessel, 2007). The ISC method cannot capture such individual differences. In considering this, initial analyses of the current string quintet data compared individual responses to acoustic changes. However, these preliminary analyses did not reveal significant individual physiological arousal responses in relation to harmonic expectancy. Therefore, the first explanation may be ruled out. A second reason could be that the effect of harmonic expectation might be more context dependent. In addition to potential differences based on listening to music in a concert context, the difference may lie also in the stimulus itself. Previous stimuli were either reduced in length (Koelsch et al., 2008; Steinbeis et al., 2006), in monophonic musical lines (Bianco et al., 2019; Egermann et al., 2013), or uniform rhythmic structures (Cheung et al., 2019). In such reduced stimuli, the harmonic surprise may have a higher element of surprise compared to naturalistic stimuli as used here, where harmony was embedded in polyphonic, full-length music. Finally, the lack of a significant effect of harmonic expectation of physiological responses could be due to the algorithm that was used to operationalise harmonic expectancy. Currently, the algorithms used to analyse harmonic expectation vary widely for accuracy of actual human perception of harmony (Harrison et al., 2020; Kern et al., 2022). Although it is generally considered that symbolic-based analysis techniques (e.g., with MIDI) are more effective in extracting harmony compared to acoustic-based analyses, the string quintet music used in Chapter 2-3 was not available in symbolic format. Computational transformation of the music to a symbolic (e.g., MIDI) format with currently available software would not be accurate as the music was polyphonic. Additionally, string instruments have a more variable onset rise, which makes note onset detection ambiguous. Harmonic expectancy in the music was operationalised here with acoustic-based methods, namely with *key clarity* from MIRToolbox (see General Methods). As this acoustic-based algorithm is considered not to perform as well as symbolic-based algorithms; harmonic expectation extracted from a different algorithm could have been a significant predictor for physiological arousal.

Although the acoustic-based features revealed that only tempo was a predictor of synchrony, the score-based analysis revealed that synchronised physiological arousal was related to a combination of higher-level features. For example, themes that re-occurred in surprising harmonic contexts (e.g., a motive that was in a major key repeated in minor key) were found to evoke synchronised physiological arousal responses. This suggests that harmony in combination with thematic and structural features likely evoke peripheral emotion responses rather than simply isolated harmonic expectancy. This aligns with studies exploring harmonic syntax. Typically, behavioural and neural indices of surprise occur in unexpected harmonies when presented at a structurally important location in the musical phrase, i.e., at the end of the musical sequence (Brattico et al., 2010; Koelsch, 2009; Koelsch et al., 2000). This effect is attenuated when the harmonic unexpected harmonic chord occurs in a less structurally important location, i.e., in the middle of the sequence (Koelsch et al., 2000). Taken together, this suggests that peripheral physiology may be influenced by harmony embedded with other features at important thematic and structural moments of a musical piece.

By going beyond the exploration of single musical features, these results demonstrated that physiological emotion was likely influenced by an interaction of various musical elements. Challenges with acoustic-based harmonic characterisation highlight the importance of developing models that quantify harmonic features of naturalistic music more precisely (Harrison et al., 2020). Additionally, results point to the importance of considering structural and thematic elements of the music on a more global level. By using both acoustic-based or score-based methods, this work illustrates how holistic musical characterisation provides a more comprehensive account of music listening experience.

2.3 Aesthetic appreciation increases with visual information

Another holistic aspect of experiencing music at a concert is being able to see the performer in conjunction with listening to the music. In laboratory studies, this visual information has been shown to increase emotional intention and expressivity (Broughton & Stevens, 2009; Dahl & Friberg, 2007; Lange et al., 2022; Luck et al., 2010; Van Zijl & Luck, 2013; Vines et al., 2004, 2011; Vuoskoski et al., 2014) and aesthetic appreciation (Platz & Kopiez, 2012). In terms of peripheral responses, previous findings from laboratory studies have been contradictory. Compared to AO music performances, an AV condition has been shown to both increased (Chapados & Levitin, 2008) or decreased (Vuoskoski et al., 2016) skin conductance (SC) arousal responses. One study by Coutinho & Scherer (2017b) explored emotional ratings between different modality conditions in a concert-like setting. Their findings were generally consistent with laboratory studies: aesthetic appreciation and wonder were highest in the audio-visual (AV) version, while boredom was highest in the audio-only (AO) version. However, Coutinho and Scherer's work was limited to Romantic Lieder[13], i.e., vocal music. Crucially, the expressive lyrics here could have had a substantial effect on emotions. In a similar way to the opera studies (see above), it is of interest to extend Coutinho and Scherer's work findings to instrumental music where the emotion is not communicated through text.

This thesis contributes to the understanding of the effect of modality in concert performances using instrumental music. In Chapter 2, the movement that was presented as an audio-only recording had significantly lower engagement to all other movements of live audio-visual performances. Assessing this more systematically, we demonstrate that audio-visual performances evoked higher self-reported aesthetic experience (Chapter 4) and engagement (Chapter 5) than to audio-only performances. Importantly, by measuring peripheral responses, we were able to gain insight into why visual information enhances aesthetic appreciation. Heart rate synchrony across participant responses was higher in audio-visual compared to audio-only music performances (Chapter 5). As previous research has shown that engagement to the stimuli is related to neural (Dmochowski et al., 2012; Kaneshiro et al., 2020) and HR synchrony (Pérez et al., 2021), this suggests that the audio-visual captured greater engagement from the audience compared to the audio-only performance. Additionally, we found partial evidence from linear mixed models that seeing the performer evoked motor-related responses of increased respiration rate and facial muscle activity (Chapter 4). Aesthetic experience was also positively associated with activation of the smiling muscle (Chapter 4). These results support the theory of embodied aesthetics. This theory suggests that seeing (implied) motion in art/performance can evoke embodied simulation of observed movements, which subsequently increases appreciation of artwork. Indeed, other research had already shown that an increase in liking is related to an increase in muscle

[13] A genre of music that typically encompassed a solo voice and piano, where text would utilise romantic poems. (Böker-Heil et al., 2001)

activity in dance (Kirsch et al., 2016) and art appreciation (Finisguerra et al., 2021). Although embodied cognition plays a large role in music psychology (Toiviainen et al., 2010), there has been a paucity of research on perceptual liking and embodied music listening (Ticini et al., 2015). Thus, the current results provide an important contribution in support of the role embodied simulation plays for aesthetic appreciation of music.

However, evidence from both Coutinho & Scherer (2017b) and the current presented studies have potential confounds. Coutinho & Scherer (2017b) incorporated four modality conditions: AV as a live performance presented in a church concert venue, and AV, AO and VO recordings that were presented in a university lecture hall. Therefore, results from Coutinho & Scherer between modality are potentially confounded with the change of context (church venue vs lecture hall). As the AV and AO conditions used here were presented in the same context, evidence is provided that an aesthetic experience may be more enhanced in an AV than an AO performance, regardless of their context. Yet, both Coutinho & Scherer (2017b) and the current studies were confounded by the live nature of the performance. In Chapter 2, 4, and 5, the AV performances were presented live, while the AO were presented as recordings. Therefore, it could be that the liveness of the performance could have mediated the emotional and peripheral responses, which may be related to the exciting spontaneity of the performance (Brown & Knox, 2017) , even if it means there is a danger of performers making mistakes (Waddell & Williamon, 2017). Some initial work in this area does show an effect of liveness on people's experience of music performance (Swarbrick et al., 2019; Tervaniemi et al., 2021), although this is not always consistent (Belfi et al., 2021). Additionally, there may have also been differences in sound quality of sound produced by live instruments vs those coming from recordings played on speakers. Therefore, further work could be done here to identify whether performer movement or liveness have a greater impact on the aesthetic experience and corresponding physiological responses.

So far, emotional and aesthetic ratings between modality in naturalistic contexts have been scarcely collected and analysed. Additionally, the comparison of physiological responses between music performances presented in different modalities has only been studied in laboratory settings. Hence, the current thesis study expands the minimal prior evidence of concert-related modality effect and breaks new ground into understanding these aesthetic responses to music performances. We conclude that seeing performer movement in a concert context has a positive contribution to the aesthetic experience. Mechanisms that might contribute to this aesthetic appreciation might arise from greater engagement with the stimulus and increased motor simulation. However, larger sample sizes would be required to further confirm the partial evidence found there. Additionally, the aspect of liveness is potentially entangled here. Therefore, further research would be required to understand the real influence of modality as separated from liveness.

2.4 Physiological measures in naturalistic musical listening

The empirical chapters showed an overall relationship with self-reports and physiological data. We replicated findings from both laboratory (e.g., Bernardi et al., 2006; Gomez & Danuser, 2004; 2007; Iwanaga, 2004; Krumhansl, 1997; Bartlett, 1996; Hodges, 1996; Hodges, 2011; 2012; Koelsch & Jäncke et al., 2015) and concert studies (Bernardi et al., 2017; Chabin et al., 2022; Egermann et al., 2013; Egermann & Reuben, 2020) that arousing music features are mirrored both in the self-reported emotion (Chapter 2, 3) and physiological responses (Chapter 2). Such mirroring of musical emotion with listeners' emotion supports previous work showing that a key mechanism of emotion induction in music is emotional contagion (Merrill, Omigie, & Wald-Furhmann, 2020). We additionally show tentative evidence that increased respiration rate and facial muscle activity predictors of aesthetic evaluations (Chapter 4). As respiration and facial muscle is innervated by the motor (somatic) system, their activation might indicate motor simulation may be a mechanism that evokes aesthetic experiences (Freedberg & Gallese, 2007).

However, there were some inconsistencies across findings. While Chapter 4 reported that the smiling muscle aligned with liking, this did not match the results presented in Chapter 2. Chapter 2 showed that liking ratings were actually associated with decreasing smile muscle activity. Indeed, some studies suggested that the activity of the zygomaticus is not necessarily specific to smiling but could rather reflect other facial expressions such as a grimace or an expression of dislike for unpleasant music (Dellacherie et al., 2011; Merrill, Ackerman, & Czepiel, 2023). One reason for this discrepancy in findings might be individual variability. In particular, expertise might play a role in motoric responses. In rating dance performances, the positive relationship of liking and smiling muscle activity was significantly stronger in dancers compared to non-dancers (Kirsch et al., 2016). When assessing physiological responses during music listening, Dellacherie et al. (2011) found that the smiling muscle increased when listening to disliked music, but significantly so in participants with higher musical expertise. Indeed, musical training can lead to changes in brain structure in (functional symmetry of) motor networks (Burunat et al., 2015; Zatorre et al., 2007), suggesting that musical expertise might indeed play an important role in individual differences of muscle responses. Futher, music preference might play a role. Indeed, when controlling for acoustic physiological arousal changed depending on listeners preference of the music in terms of liking (Salimpoor et al., 2009) and disliking (Merrill, Ackermann, & Czepiel, 2023). Although it was not within the scope of this thesis to consider individual differences (such as musical expertise or preferences) across our samples, such comparisons would be crucial to further understand how reliable or flexible peripheral responses are across individuals.

From investigating peripheral measures, I conclude that peripheral response signatures can be related to a) acoustic-induced emotion potentially induced by means of emotional contagion (Juslin, 2016) and b) subjective states such as liking, likely through mechanisms of embodied appreciation (Freedberg & Gallese, 2007). Peripheral responses reflecting arousal seemed stable across all empirical studies. However, the relationship between liking and the zygomaticus ('smiling') muscle were not consistent across findings. From this evidence, we cannot conclude that peripheral responses yield a reliable indicator of aesthetic experience per se. Responses may be inconsistent due to individual differences of musical expertise and preferences. Therefore, rather than indexing aesthetic experiences directly, it seems peripheral measures may be a useful complement to understanding (mechanisms of) emotional and aesthetic responses.

2.4 Peripheral synchrony: new insights into naturalistic musical listening

In the field of auditory neuroscience, one way to assess experience of naturalistic stimuli involves calculating synchrony of responses between participants (Dmochowski et al., 2012; Hasson et al., 2004; 2010; Kaneshiro et al., 2019; Lillywhite et al., 2022; Madsen et al., 2019; Nastase et al., 2019). To date, this was mainly explored in neural responses with fMRI (e.g., Hasson et al., 2004; Lillywhite et al., 2022) and EEG (e.g., Dmochowski et al., 2012; Kaneshiro et al., 2019; Madsen et al., 2019). However, there has been growing interest to further investigate peripheral responses to gain insight into a more holistic representation of human perception and cognition (Azzalini et al., 2019; Criscuolo et al., 2022; Klimesch, 2018; Madsen et al., 2022; Parviainen, Lyyra, & Nokia, 2022). Studies are beginning to extend work on synchrony measures to peripheral responses such as skin conductance (Bracken et al., 2014), heart rate (Perez et al., 2021), respiration rate, and eye activity (Madsen et al., 2022). The current thesis contributes to this growing body of research in peripheral responses. Inter-subject synchrony was originally used in Chapter 3 to assess systematic responses with high signal-to-noise ratio (Hasson et al., 2004; Nastase et al., 2019). Chapter 5 builds on this, providing a more comprehensive investigation of physiological synchrony. Here, related measures of stimulus-response synchrony were also assessed, as was inter-subject and stimulus-response synchrony in the phase domain.

By using peripheral responses, greater insights could be achieved into what responses are being synchronised. Most work on time-resolved music listening has used EEG. To calculate ISC with EEG, the multivariate data (i.e., responses across several EEG channels) typically undergoes dimension reduction. Reduction techniques collapse several channels of data to fewer channels, while maximising variance in the data. ISC studies have done this either with canonical component analysis (CCA) or reliable components analysis (RCA). The topographies of the extracted components seem to be similar across both musical (Kaneshiro et al., 2020; Madsen et al., 2019) and non-musical auditory natural stimuli (Ki et al., 2016), showing symmetric parietal negativity and occipital positivity. This component is broadly interpreted to reflect the processing of temporally structured auditory stimuli. In using peripheral responses (which are univariate), it was not necessary to reduce the data to calculate synchrony. Therefore, it was possible to assess more precisely what response is synchronised. In Chapter 3, ISC was shown to be a successful method in finding moments where participants had a typical arousal response, i.e., increased of RR and SC reflecting activation of a sympathetic nervous system (see Figure 1 in Chapter 3). High synchrony moments also indicated a HR increase, though this increase was not significant after Bonferroni adjustments. In Chapter 5, high ISC was related to deceleration-acceleration patterns in the HR and RR (see Supplementary Materials, Chapter 5). These responses are thought to reflect increase in sympathetic activity (Chapter 3, 5), i.e., emotional arousal, as well as orienting responses (Chapter 5), i.e., enhanced attention to a stimulus. Hence, in extending the idea that neural synchrony is related to auditory processing (Kaneshiro et al., 2020; Ki et al., 2016; Madsen et al., 2019), these current results show that peripheral synchrony is related to physiological indices of higher arousal and attention.

To the best of my knowledge, Chapters 3 and 5 are one of the first contributions into the field that explores peripheral synchrony in a naturalistic concert context (see also Chabin et al., 2022; Tschacher …, Czepiel, Tröndle, & Meier, 2021). As has been highlighted throughout this thesis, assessing whether findings from laboratory studies are stable in more real-world settings is becoming increasingly important (Tervaniemi, 2023). The replication of results across settings would suggest robustness of effects, while differences between them would suggest that music listening experiences are context dependent. We were able to replicate the effect that physiological synchrony was related to emotional engagement and to structural events in music (Chapters 3 and 5) (Dauer et al., 2021; Kaneshiro et al., 2020). Such replication suggests that these responses might be robust across laboratory and real-world settings. However, Madsen & Parra (2022) did not find respiration rate to be significantly synchronised across participants; but we do (Chapter 5). In assuming that synchrony reflects engagement, it could be interpreted that music within a concert setting (Chapter 5) is more engaging than instructional videos presented in a laboratory setting (Madsen & Parra, 2022). This is an example where the results are context-dependent, highlighting the importance of assessing such engagement in a setting that allows genuine absorption and more intense emotional responses (Lamont, 2011).

However, there are differences between time- and phase-domain synchrony. As Chapter 3 speculated that physiological synchrony could be related to phase-reset (Peelle & Davis, 2012; Schroeder & Lakatos, 2009; Voloh & Womelsdorf, 2016), Chapter 5 explored not only time-domain synchrony (i.e., correlation), but additionally phase-domain synchrony (i.e., coherence). We replicated laboratory studies (Madsen et al., 2019) that time-domain synchrony was overall significant during music listening compared to control (circular shifted) data (Chapter 5). Regarding phase-domain synchrony, while laboratory studies have found that breathing aligns with musical beats (Etzel et al., 2006; Haas et al., 1986), Chapter 5 did not corroborate this finding in a concert setting. Although counter-intuitive, it was expected that more engagement with stimuli in a concert setting would increase chances of phase synchrony. However, it can also be explained by the fact that the musical beat is more prominent if heard in an isolated laboratory setting. Within a concert setting, there are several other factors that contribute to the experience. With an increase of other inputs, perhaps the low-level musical beat seems less important, but greater importance is placed on higher-level features, such as structure (Chapter 2, 3, 5, see section 2.1). This also highlights the differences between time-domain and phase-domain synchrony; while time domain synchrony might be more related to engagement (Pérez et al., 2021), phase domain synchrony might not be. Phase synchrony might rather be related to the stimulus intelligibility (Harding et al., 2019; Luo & Poeppel, 2007; Peelle et al., 2013; Vanden Bosch der Nederlanden et al., 2020). Additionally, phase synchrony might be different depending on whether participants can move. Phase synchrony occurs when participants can move more freely and interact with each other (Gugnowska et al., 2022; Hartmann et al., 2019), which could lead to better prediction of upcoming beats (Criscuolo et al., 2022; Koelsch et al., 2019). The reason we do not replicate such results here might be due to the fact that typical Western classical concert etiquette requires audience members to limit movements and interactions. Thus, phase synchrony might become more relevant in contexts where listeners can move and interact freely, for example in pop and rock concerts, as opposed to classical music concerts.

3 Limitations and outlook

3.1 The effect of the concert context

This thesis explored an individual's and an audience's concert experiences. The current results allow concluding that concert components of full-length music and its programming as well as the additional visual input contribute to what might make a concert special. However, in the research presented here, concert contexts were not directly compared to non-concert contexts. Thus, it is difficult to specify the effect of the concert context itself as a physical and social environment. Therefore, one unexplored aspect is to compare a concert experience to a different situational context. One particular direction that was planned, but due to COVID-19 could not be implemented, was to better understand what might cause peripheral physiological synchrony. Although ISC highlights commonalities across participants in stimulus processing (Hasson et al., 2004; Nastase et al., 2019), it could be that the ISC is also driven by the shared experience of like-minded people attending a concert. This question is motivated by prior results suggesting that physiological synchrony between participants reflects social aspects (Konvalinka et al., 2011), and is higher when participants are co-present (Golland et al., 2015). Indeed, concerts likely create a social experience (Dotov & Trainor, 2021; O'Neill & Egermann, 2022). Thus, the social presence of others might be reflected in similarity of self-reports and neural measures (Chabin et al., 2022; Golland et al., 2015). However, a comparison with co-present and non-co-present audience members could not be undertaken given the nature of the experimental design. The string quintet concerts had live performances every evening, making each performance slightly different. Even when time warping the data (which was attempted based on Thompson & Luck, 2012; Wanderley et al., 2005), results were not conclusive as to whether any differences might be related to actual co-presence or the slight deviations in the live performances. Using a recording would have been preferred. Thus, it was also considered to carry out such an analysis with the recorded AO versions from the piano concerts. However, this recorded version was also not comparable because the piece orders and modality orders between concerts were not identical. Therefore, any reported difference would not allow disentangling the effect of piece order and co-presence. In the future, experimental designs exploring the effect of co-presence could consider using a (video) recording that is stable across (groups of) participants. Such an exploration could provide valuable insights into understanding whether synchrony simply reflects commonalities in stimulus-related evoked responses (Hasson et al., 2004) or whether it could allow assessing a shared concert experience (Chabin et al., 2022; Golland et al., 2015; O'Neill & Egermann, 2022).

3.2 Generalisability

Although the current results have high ecological validity as they were collected in a typical chamber concert hall setting, the current findings are limited to the ArtLab concert hall (see General Methods). Concert halls vary substantially in features such as size, architecture styles, and acoustic properties (Lokki et al., 2012). These concert features differences might also affect any music listening experience. For example, in terms of size, the ArtLab provided a more intimate setting that was more suitable for the chamber music concerts used here. On the other hand, a much larger concert hall might evoke a greater sense of awe (Konečni, 2005). Thus, the extent to which these results are generalisable is limited to the concert hall context that was used here.

Only three Western classical music pieces were used in each concert data set. This was a representative number of pieces for a concert program and also provided a variety of known/lesser-known pieces. However, this is a much smaller number of pieces or clips that are used in other laboratory studies. Therefore, it could be argued that the current findings are ultimately specific to these musical pieces. However, adding further musical pieces would have made the concert too long. The pieces presented were appropriately typical for a (classical) concert experience and the results therefore provide a basis for future concert work. Future complementary studies could take a wider range of musical pieces into consideration, potentially at the cost of shorter pieces.

Related to limited generalisability of the pieces is that the present findings mainly result from Western classical tonal music only. The effect of emotion trajectories and tempo as a significant predictor of synchronised arousal responses only showed in Beethoven and Brahms (CPP pieces) and less in the contemporary (atonal) piece (Chapter 2-3). Piece differences were not explored in Chapters 4 and 5 as there was only one movement in each piece. Thus, there might not have been sufficient power in the data to explore piece differences. Nonetheless, the exploration and comparison of CCP and of contemporary music is an interesting venture for music neuroaesthetics. This is because much of the theories of musical emotions and pleasure are related to expectation and predictability in music based on traditional tonal and rhythmic systems. The idea that expectation leads to pleasure holds true for much of tonal music, where certainty and predictions can be made within a relatively stable harmonic framework (Cheung et al., 2019). However, this theory might not hold for atonal music, where there is no clear tonal centre (Brattico, 2021; Mencke et al., 2019). However, people enjoy hearing such complex music, suggesting that current theory does not explain other musical genres. Some research is beginning to deepen our understanding of the differences between the two genres. Tonal and atonal music have processing differences, as shown in the neural responses of MEG (Fernández-Rubio et al., 2022; Mencke et al., 2021). The liking of such music might also result from other factors such as openness to experience novel stimuli (Brattico, 2021; Mencke et al., 2019). A key aspect mentioned in these studies is that such music might be more enjoyable within an artistic context, such as a concert. Indeed, art objects are likely evaluated with higher aesthetic value if in an aesthetic context, such as a museum or concert (Leder et al., 2004; Leder & Nadal, 2014). A concert context is likely to enhance interest in and appreciation of music, especially as the focus lies on the music performance itself. Such parameters make the concert a meaningful context to explore contemporary classical music. Overall, as our findings were more related to the Western classical tonal music, future work could provide additional insights into different mechanisms of aesthetic experiences in other genres, such as contemporary classical music.

While not a limitation per se, of note is that the current findings relate to a specific participant population. The participants were recruited via concert advertisement, which is typical for the majority of classical concerts. This limits the current results to individuals who regularly attend concerts and who have at least a few years of musical experience. This sampling strategy benefitted the notion that participants were likely to attend and enjoy the music. The audience were probably familiar with at least some of the compositions that it encountered and possibly open to hearing works of lesser-known music composers. However, the audiences represent a relatively small (convenience) sample of the general population. Results could possibly differ with participants who do not regularly attend or have never attended a classical concert, simply because they may find it boring (Dearn & Pitts, 2017; Kolb, 2000). Although this was not within the scope of this thesis, this could become essential research to further clarify why people do or do not attend classical music concerts. As individual context – such as musical taste preference and musical experience – is a key input of an aesthetic experience (Brattico, 2021; Wald-Fuhrmann et al., 2021), future research could use between-group experimental designs to investigate such differences. One direction, for example, could be assessing differences between participants with varying musical expertise (see also discussion of individual differences in physiological responses, Section 2.4). Additionally, broader and larger sample size could be beneficial for potential statistical modelling.

3.3 Stimuli

On the spectrum of control and reality, the choice of the music itself was more on the side of reality. The music was chosen to represent a typical classical chamber concert, with a variety of different genres and emotions. However, specific aspects of the musical stimuli brought on certain challenges.

First, the length of the musical pieces varied. Across piano and string quintet concerts, all movements were performed in their entirety, and lengths ranged from 2.30 to 10.30 minutes. This has the advantage of completeness and is expected in real-world music listening. However, different piece lengths presented challenges. In terms of statistical analysis, an ideal control condition to assess significance would have needed pieces of the same length (Chapter 5). In cases of naturalistic paradigms where there is no control included as part of the design, studies have created a control data set by shuffling original responses from participants (e.g., Fink et al., 2023; Harding et al., 2019; Madsen et al., 2019). The assumption here is that the shuffled data would have no effect, while the original data should have an effect. Responses can be shuffled in terms of labels or data time series. In label shuffling, one would randomly shuffle data to reassign some responses from one condition to a different condition. Shuffled time series can be created by randomly shuffling data to random time points, potentially while retaining certain spectral and phase properties. Time series data can also be (circular) shifted, where time series are moved along the time axis, so responses are no longer time-locked to the start of the trial. The ideal way to create control data is to shuffle data labels, as it keeps the data intact, rather than distorting (spectral features of) the signal (Fink et al., 2023). However, to use this technique, all trials should be the same length. In the current data sets, all 'trials' (pieces) were of different lengths, thus did not allow assessing differences with a label-shuffled control. Here, rather than shuffling the labels, we shifted the data across time to create a time-'unlocked' version of peripheral responses (Chapter 5). Circular shifting was employed as it distorts the data the least (Fink et al., 2023). It is also not optimal to cut full-length music, as such cutting of music may distort harmonic and structural completeness. Future research might choose musical pieces that are more similar or even identical in length to allow for more consistent comparisons. Of note is that the pieces should perhaps be of similar content and take similar energy to listen to; a point discussed in greater detail in the following paragraph.

When comparing differences between pieces in Chapters 2 and 3, there was a potential confound of length. As mentioned above, classical and romantic pieces had higher positive ratings compared to the contemporary piece. It is most likely that these ratings were lower than in the classical contemporary piece due to its less conventional metre and harmony. However, a potential confound is that the contemporary piece was shorter. This means that the length may have led to a less positive aesthetic experience, for example, allowing less time to become absorbed in the music. Nonetheless, while the lengths differed technically, they might have equated to a similar experience perceptually and musically. This is related to the complexity of the contemporary pieces and the effort required to listen to it. Compared to regular harmony and rhythmic structures of the CPP pieces, listening to more dissonances, irregular rhythms, and ambiguous harmony in the contemporary piece might take significantly more effort to process within a shorter time span. Indeed, the choice of a shorter contemporary piece was likely an artistic consideration; many concert programmers likely would not typically want to over-balance a concert toward modern music[14]. Therefore, despite the actual lengths, in terms of cognitive load and listening effort, pieces could have been comparable.

From an experimental point of view, obtaining ratings and physiological responses for each piece limited the number of observations. This limits the power of statistical tests. Here, this was circumvented by taking advantage of the continuous recording of physiology. Either a continuous analysis method was employed (Chapter 3) or physiological data across entire music pieces were cut into structural sections to create more ‘trials’ (Chapter 4). Therefore, although exploring responses to such long pieces can be beneficial in gaining new insights, complementary future research should rely on more participants, shorter pieces, or develop new methodological ways to boost statistical power.

The variety of music features in naturalistic stimuli is limited to the variation in the music itself. Some research has considered this by choosing a piece that varied greatly in timbral, rhythmic, and harmonic features (e.g., Piazolla’s *Adios Nonino* in Alluri et al., 2012; Burunat et al., 2014). Here, the feature range was distributed relatively evenly across the twelve movements used in the string quintet. However, the piano concert music pieces varied less in the tempi and rhythmic and harmonic features, limiting statistical analyses and possible conclusions regarding acoustic features (Chapters 4 and 5). Similarly in the string quintet, tempo was predictive of physiological synchrony, but not for the contemporary piece of music (Chapter 3). As the tempo variance was different across the pieces, it might be that this did not allow for a fair statistical comparison. However, this is a needed compromise when using naturalistic

[14] There are exceptions, for example if certain festivals are more oriented to contemporary classical period music, such as the Aldeburgh Music Festival, U.K. or the Darmstadt International Summer Courses for New Music, Darmstadt, Germany.

music. Nonetheless, future research should aim to allow enough variance in acoustic features for fair and appropriate statistical comparisons, especially across pieces or genres.

An advantage across all studies was that musical pieces were performed live, which likely enhanced the audience's engagement. However, live music performances differ across repetitions, even when performances are highly polished (Chaffin et al., 2007). This was minimised by hiring professional performers who were instructed to play as similarly as possible across all music performances. Acoustic features were compared to confirm the similarity across performances. Nonetheless, there were slight deviations between live performances, such as tempo changes, occasional mistakes, variant loudness and timbre. While it is not possible to combat differences of certain features, such as mistakes, there are ways to combat other features. For example, it is possible to align timing. Alignment was essential here to allow comparison of continuous responses across concerts, particularly for the results presented in Chapter 2. Time alignment requires timestamps (i.e., triggers) across performances. As live performances are spontaneous, adding timestamps (triggers) cannot be planned ahead. Therefore, timestamps must be manually annotated either online or offline. Once timestamps have been added, data can be aligned based on these time stamps. Previous work has used dynamic time warping algorithms (Thompson & Luck, 2012; Wanderley et al., 2005). However, this computationally is expensive and might introduce distortions to the data. Therefore, to align physiological data in Chapter 2, we opted to parse data in meaningful units (Kang & Wheatley, 2017), bars (American: measures), and average over these units. Overall, live performances are much more realistic in terms of real-world music listening, sometimes at the cost of inevitable inconsistencies that cannot be controlled or diminished, e.g., performer mistakes. Nonetheless there are possibilities to correct the deviations (e.g., timing deviations) offline.

Despite having audio-visual performances, we mostly focused on the auditory aspects. It was not feasible to explore the visual features. Although it is possible to extract movements from video recordings (Camurri et al., 2004; Jakubowski et al., 2017), the angle of the video recording taken from the concerts was not ideal for analysing the movements of the piano player nor the string instrumentalists. We did not have a camera at an angle that would capture the musicians from the front, but an angle that peered over the right shoulder of the piano player (see Chapter 4, Figure 1) and the first violinist. This meant that the back of the piano player and the first violinist was facing the camera, which blocked most of the view of their left hand and arm. Therefore, future studies using a visual component should consider testing for a good video recording at an appropriate angle (or even multiple cameras for better coverage) to capture the majority movements.

Finally, the visual effect reported in Chapters 4 and 5 was dependent on the musicians themselves. Some performers move a lot more compared to other performers. These different performer movement styles were shown to influence participants' ratings (Behne & Wöllner, 2011; Griffiths, 2010). Therefore, although we corroborated previous findings (Platz & Kopiez, 2012), findings from Chapters 4-5 might be restricted to the performers' movement style.

3.4 Physiology

Physiological responses were related to both acoustic features of music and subjective states of individual listeners. Rather than substituting self-report measures, peripheral responses could be seen as complementing self-report measures. However, it still seems vital that self-reports are collected alongside physiology to understand and interpret them appropriately. Another potential limitation in interpreting physiological results is the adequate comparison to a traditional, neutral baseline. During data collection for the string quintet and piano concert, three minutes 'baseline' recordings for the peripheral physiological responses were actually collected. However, it was decided against using this data for three main reasons. First, the baseline was not meaningful as it was not long enough to compare the pieces. Second, these 'baseline' recordings were collected right before a concert. Measuring a baseline before a laboratory experiment may be an appropriate test of a 'resting state' physiology. However, measuring physiology right before a concert could have captured the excitement of anticipating live music performances. Therefore, the baseline collected does not reflect a true 'resting' state. Third, the baseline recording was collected when the musicians were already on the stage. According to the videos of the baseline recording, the musicians fidgeted slightly and looked around the room. Thus, the responses from the audience were not neutral, but time locked to these movements. Other common events such as someone in the audience coughing also added to the issue. Indeed, it is a challenge to sometime find an adequate control for naturalistic stimuli, as real-world experience does not necessarily always entail comparisons and contrasts (Schön & Morillon, 2019; Tervaniemi, 2023). To circumvent this lack of baseline in the design, we followed previous research using naturalistic stimuli, by creating baselines for comparison from the data itself: moments of low synchrony are compared to moments of high synchrony (Chapter 3) and time-unlocked (circular shifted) control data sets are compared to true datasets (Chapter 5). Nonetheless, actual baselines and other comparative conditions incorporated in the designs might be more accurate to further elucidating what drives physiological response. For example, comparing synchrony between audiences experiencing a music performance and audiences sitting in silence with an empty stage might shed light on whether synchrony is being driven more by the music or by the co-presence of other audience members. Thus, to further understand physiological responses in ecologically valid settings, future studies would benefit not only from additional self-reports, but also experimental baselines/comparative conditions to help interpret physiological responses.

The synchrony measures also revealed some limitations. ISC enhances the commonalities across a signal, while filtering out idiosyncratic responses to stimuli (i.e., individual responses) and spontaneous activity. However, it might prove insightful to explore individual responses, especially for aesthetic responses. Studies show that aesthetic responses might deviate between individuals in the visual domain (Vessel & Rubin, 2010). In the music domain, there are more similarities across individuals in terms of cognitive processing, but more deviations in aesthetic judgements such as beauty (Müller et al., 2009). Consequently, other methods might still prove useful in assessing peripheral and neural responses in naturalistic stimuli. Indeed, there are several exciting new methods that are being developed to assess responses to naturalistic stimuli, such as mutual information (Chalas et al., 2023) and temporal-response-functions (Di Liberto et al., 2020; Kern et al., 2022; Weineck et al., 2022). Nonetheless, the synchrony methods used here proved to be a useful method in providing initial steps in assessing peripheral responses in naturalistic stimuli.

4 Conclusion

By applying techniques from laboratory studies as well as developing novel and exploratory methods, this thesis aimed to gain insight into real-world music listening in a concert setting. The development of more portable and wearable technologies leads us beyond the laboratory to more naturalistic settings where music can be experienced in more immersive situations. By conducting studies in ecological contexts, this thesis extends our understanding of music listening and highlights that such research in more challenging environments is not only possible, but worthwhile and interesting. Although the results might not have as clear conclusions as certain laboratory and controlled experimental designs, the main interest is the validity of the current results. The current thesis highlights how emotions and engagement fluctuate over an entire music concert program, over multimovement works, and at local structural changes in the music. Such contrasting emotions and engagement evoked by different music styles and moods may enhance interest for listeners. Perceiving structural changes might aid in chunking the music to facilitates processing, which in turn might enhance aesthetic experience. I also show how aesthetic experience and peripheral responses are affected by a holistic interplay of music and visual features. By exploring (synchrony of) peripheral physiological responses, mechanisms that might contribute to this aesthetic appreciation of music might arise from emotional contagion (when the emotion of the music is mirrored in the emotion of the listener), greater engagement with the stimulus, and increased motor simulation. Although some results regarding peripheral responses were consistent across studies, some were not. Differences arose in the zygomaticus ('smiling') muscle and respiration. As these are related to the motor system, and audio-motor differences have been found to vary between musical expertise, these discrepancies in findings point to potential individual differences based on expertise and preferences. Currently these findings do not fully allow concluding that peripheral physiological responses can (yet) substitute self-reports of aesthetic experience. Nonetheless, peripheral responses were shown to complement self-reports and provided crucial insights into the potential mechanisms that might drive emotional and aesthetic appreciation.

In addition to corroborating and extending previous work, open questions emerged. Future directions can build on this research by comparing different concert and situational contexts, assessing whether synchrony purely reflects stimulus tracking or whether it also reflects social co-presence. This thesis also advocates the careful consideration of experimental as well as artistic aspects in future research of this type. Despite the identified challenges, testing music experiences in real life settings is an important step forward in the research of music aesthetics. Although more complex and less simplistically clear, the findings have more relevance in terms of understanding real-world listening to music. Overall, these findings contribute to the growing body of studies exploring not only naturalistic music listening experiences, but also real-world music listening in naturalistic contexts.

References

Abrams, D. A., Ryali, S., Chen, T., Chordia, P., Khouzam, A., Levitin, D. J., & Menon, V. (2013). Inter-subject synchronization of brain responses during natural music listening. *European Journal of Neuroscience*, *37*(9), 1458–1469. https://doi.org/10.1111/ejn.12173

Adams, K., Springer, D. G., Geringer, J. M., Nápoles, J., & Davis, M. N. (2023). Effects of Transitional Elements on Listeners' Aesthetic Responses to a Choral Performance. *Journal of Research in Music Education*, *71*(1), 111–128. https://doi.org/10.1177/00224294221111715

Al, E., Iliopoulos, F., Forschack, N., Nierhaus, T., Grund, M., Motyka, P., Gaebler, M., Nikulin, V. V., & Villringer, A. (2020). Heart–brain interactions shape somatosensory perception and evoked potentials. *Proceedings of the National Academy of Sciences*, *117*(19), 10575–10584. https://doi.org/10.1073/pnas.1915629117

Albouy, P., Benjamin, L., Morillon, B., & Zatorre, R. J. (2020). Distinct sensitivity to spectrotemporal modulation supports brain asymmetry for speech and melody. *Science*, *367*(6481), 1043–1047. https://doi.org/10.1126/science.aaz3468

Alday, P. M. (2019). M/EEG analysis of naturalistic stories: A review from speech to language processing. *Language, Cognition and Neuroscience*, *34*(4), 457–473. https://doi.org/10.1080/23273798.2018.1546882

Alluri, V., & Toiviainen, P. (2010). Exploring Perceptual and Acoustical Correlates of Polyphonic Timbre. *Music Perception*, *27*(3), 223–242. https://doi.org/10.1525/mp.2010.27.3.223

Alluri, V., Toiviainen, P., Jääskeläinen, I. P., Glerean, E., Sams, M., & Brattico, E. (2012). Large-scale brain networks emerge from dynamic processing of musical timbre, key and rhythm. *NeuroImage*, *59*(4), 3677–3689. https://doi.org/10.1016/j.neuroimage.2011.11.019

Alluri, V., Toiviainen, P., Lund, T. E., Wallentin, M., Vuust, P., Nandi, A. K., Ristaniemi, T., & Brattico, E. (2013). From Vivaldi to Beatles and back: Predicting lateralized brain responses to music. *NeuroImage*, *83*, 627–636. https://doi.org/10.1016/j.neuroimage.2013.06.064

Ames, D. L., Honey, C. J., Chow, M. A., Todorov, A., & Hasson, U. (2015). Contextual Alignment of Cognitive and Neural Dynamics. *Journal of Cognitive Neuroscience*, *27*(4), 655–664. https://doi.org/10.1162/jocn_a_00728

Anglada-Tort, M., & Skov, M. (2020). What Counts as Aesthetics in Science? A Bibliometric Analysis and Visualization of the Scientific Literature From 1970 to 2018. *Psychology of Aesthetics, Creativity, and the Arts*, *16*(3), 553–568.

Anglada-Tort, M., Thueringer, H., & Omigie, D. (2019). The busking experiment: A field study measuring behavioral responses to street music performances. *Psychomusicology: Music, Mind, and Brain*, *29*(1), 46–55. https://doi.org/10.1037/pmu0000236

Ara, A., & Marco-Pallarés, J. (2020). Fronto-temporal theta phase-synchronization underlies music-evoked pleasantness. *NeuroImage*, *212*, 116665. https://doi.org/10.1016/j.neuroimage.2020.116665

Ara, A., & Marco-Pallarés, J. (2021). Different theta connectivity patterns underlie pleasantness evoked by familiar and unfamiliar music. *Scientific Reports*, *11*(1), 18523. https://doi.org/10.1038/s41598-021-98033-5

Ardizzi, M., Calbi, M., Tavaglione, S., Umiltà, M. A., & Gallese, V. (2020). Audience spontaneous entrainment during the collective enjoyment of live performances: Physiological and behavioral measurements. *Scientific Reports*, *10*(1), 3813. https://doi.org/10.1038/s41598-020-60832-7

Arnal, L. H., Morillon, B., Kell, C. A., & Giraud, A.-L. (2009). Dual Neural Routing of Visual Facilitation in Speech Processing. *The Journal of Neuroscience*, *29*(43), 13445–13453. https://doi.org/10.1523/JNEUROSCI.3194-09.2009

Arnqvist, G. (2020). Mixed Models Offer No Freedom from Degrees of Freedom. *Trends in Ecology & Evolution*, *35*(4), 329–335. https://doi.org/10.1016/j.tree.2019.12.004

Assaneo, M. F., Rimmele, J. M., Sanz Perl, Y., & Poeppel, D. (2021). Speaking rhythmically can shape hearing. *Nature Human Behaviour*, *5*(1), 71–82. https://doi.org/10.1038/s41562-020-00962-0

Auslander, P. (2002). *Liveness: Performance in a Mediatized Culture* (1 ed.). Routledge. https://doi.org/10.4324/9780203006955

Azzalini, D., Rebollo, I., & Tallon-Baudry, C. (2019). Visceral Signals Shape Brain Dynamics and Cognition. *Trends in Cognitive Sciences*, *23*(6), 488–509. https://doi.org/10.1016/j.tics.2019.03.007

Bach, D. R., Neuhoff, J. G., Perrig, W., & Seifritz, E. (2009). Looming sounds as warning signals: The function of motion cues. *International Journal of Psychophysiology*, *74*(1), 28–33. https://doi.org/10.1016/j.ijpsycho.2009.06.004

Bacha-Trams, M., Ryyppö, E., Glerean, E., Sams, M., & Jääskeläinen, I. P. (2020). Social perspective-taking shapes brain hemodynamic activity and eye movements during movie viewing. *Social Cognitive and Affective Neuroscience*, *15*(2), 175–191. https://doi.org/10.1093/scan/nsaa033

Balteș, F. R., & Miu, A. C. (2014). Emotions during live music performance: Links with individual differences in empathy, visual imagery, and mood. *Psychomusicology: Music, Mind, and Brain*, *24*(1), 58–65. https://doi.org/10.1037/pmu0000030

Bangert, M., Peschel, T., Schlaug, G., Rotte, M., Drescher, D., Hinrichs, H., Heinze, H.-J., & Altenmüller, E. (2006). Shared networks for auditory and motor processing in professional pianists: Evidence from fMRI conjunction. *NeuroImage*, *30*(3), 917–926. https://doi.org/10.1016/j.neuroimage.2005.10.044

Bannister, S., & Eerola, T. (2018). Suppressing the Chills: Effects of Musical Manipulation on the Chills Response. *Frontiers in Psychology*, *9*, 2046. https://doi.org/10.3389/fpsyg.2018.02046

Barr, D. J. (2021). *Learning statistical models through simulation in R: An interactive textbook* (Version 1.0.0). https://psyteachr.github.io/stat-models-v1

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Barrett, L. F., & Russell, J. A. (2015). *The Psychological Construct of Emotion*. The Guilford Press.

Barry, R. J. (1975). Low-intensity auditory stimulation and the GSR orienting response. *Physiological Psychology*, *3*(1), 98–100. https://doi.org/10.3758/BF03326832

Barry, R. J., & Sokolov, E. N. (1993). Habituation of phasic and tonic components of the orienting reflex. *International Journal of Psychophysiology*, *15*(1), 39–42. https://doi.org/10.1016/0167-8760(93)90093-5

Bartlett, D. L. (1996). Physiological responses to music and sound stimuli. In *Handbook of Music Psychology* (pp. 343–385).

Barton, K. (2015). *Mu-MIn: Multi-model inference. R Package* (1.43.17). http://R-Forge.R-project.org/projects/mumin/

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using {lme4}. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Baumgarten, A. G. (1735). *Reflections on poetry: Alexander Gottlieb Baumgarten's Meditations philosophicae de nonnullis ad poema pertinentibus*. University of California Press.

Baumgartner, T., Esslen, M., & Jäncke, L. (2006). From emotion perception to emotion experience: Emotions evoked by pictures and classical music. *International Journal of Psychophysiology*, *60*(1), 34–43. https://doi.org/10.1016/j.ijpsycho.2005.04.007

Behne, K.-E., & Wöllner, C. (2011). Seeing or hearing the pianists? A synopsis of an early audiovisual perception experiment and a replication. *Musicae Scientiae*, *15*(3), 324–342. https://doi.org/10.1177/1029864911410955

Belfi, A. M., Samson, D. W., Crane, J., & Schmidt, N. L. (2021). Aesthetic Judgments of Live and Recorded Music: Effects of Congruence Between Musical Artist and Piece. *Frontiers in Psychology*, *12*. https://www.frontiersin.org/article/10.3389/fpsyg.2021.618025

Benedek, M., & Kaernbach, C. (2010). A continuous measure of phasic electrodermal activity. *Journal of Neuroscience Methods*, *190*(1), 80–91. https://doi.org/10.1016/j.jneumeth.2010.04.028

Berlyne, D. E. (1971). *Aesthetics and psychobiology*. Appleton-Century-Crofts.

Bernardi, L., Porta, C., & Sleight, P. (2006). Cardiovascular, cerebrovascular, and respiratory changes induced by different types of music in musicians and non-musicians: The importance of silence. *Heart*, *92*(4), 445–452. https://doi.org/10.1136/hrt.2005.064600

Bernardi, N. F., Codrons, E., di Leo, R., Vandoni, M., Cavallaro, F., Vita, G., & Bernardi, L. (2017). Increase in Synchronization of Autonomic Rhythms between Individuals When Listening to Music. *Frontiers in Physiology*, *8*, 785. https://doi.org/10.3389/fphys.2017.00785

Berntson, G. G., Quigley, K. S., Norman, G. J., & Lozano, D. L. (2016). Cardiovascular Psychophysiology. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of Psychophysiology* (4th ed., pp. 183–216). Cambridge University Press. https://doi.org/10.1017/9781107415782.009

Bevilacqua, D., Davidesco, I., Wan, L., Chaloner, K., Rowland, J., Ding, M., Poeppel, D., & Dikker, S. (2019). Brain-to-Brain Synchrony and Learning Outcomes Vary by Student–Teacher Dynamics: Evidence from a Real-world Classroom Electroencephalography Study. *Journal of Cognitive Neuroscience*, *31*(3), 401–411. https://doi.org/10.1162/jocn_a_01274

Bianco, R., Gold, B. P., Johnson, A. P., & Penhune, V. B. (2019). Music predictability and liking enhance pupil dilation and promote motor learning in non-musicians. *Scientific Reports*, *9*(1), 17060. https://doi.org/10.1038/s41598-019-53510-w

Bianco, R., Novembre, G., Keller, P. E., Villringer, A., & Sammler, D. (2018). Musical genre-dependent behavioural and EEG signatures of action planning. A comparison between classical and jazz pianists. *NeuroImage*, *169*, 383–394. https://doi.org/10.1016/j.neuroimage.2017.12.058

Bigand, E., & Poulin-Charronnat, B. (2006). Are we “experienced listeners”? A review of the musical capacities that do not depend on formal musical training. *Cognition*, *100*(1), 100–130. https://doi.org/10.1016/j.cognition.2005.11.007

Blackburn, A. (2016). The role repertoire choice has in shaping the identity and functionality of a chamber music ensemble. *ICMPC14*, 583–587.

Blood, A. J., & Zatorre, R. J. (2001). Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proceedings of the National Academy of Sciences*, *98*(20), 11818–11823. https://doi.org/10.1073/pnas.191355898

Blood, A. J., Zatorre, R. J., Bermudez, P., & Evans, A. C. (1999). Emotional responses to pleasant and unpleasant music correlate with activity in paralimbic brain regions. *Nature Neuroscience*, *2*(4), 382–387. https://doi.org/10.1038/7299

Boiten, F. A., Frijda, N. H., & Wientjes, C. J. E. (1994). Emotions and respiratory patterns: Review and critical analysis. *International Journal of Psychophysiology*, *17*(2), 103–128. https://doi.org/10.1016/0167-8760(94)90027-2

Böker-Heil, N., Fallows, D., Baron, J. H., Parsons, J., Sams, E., Johnson, G., & Griffiths, P. (2001). Lied. In *Oxford Music Online*. Oxford University Press. https://doi.org/10.1093/gmo/9781561592630.article.16611

Bonetti, L., Brattico, E., Vuust, P., Kliuchko, M., & Saarikallio, S. (2020). Intelligence and Music: Lower Intelligent Quotient Is Associated With Higher Use of Music for Experiencing Strong Sensations. *Empirical Studies of the Arts*, 027623742095141. https://doi.org/10.1177/0276237420951414

Bracken, B. K., Alexander, V., Zak, P. J., Romero, V., & Barraza, J. A. (2014). Physiological Synchronization Is Associated with Narrative Emotionality and Subsequent Behavioral Response. In D. D. Schmorrow & C. M. Fidopiastis (Eds.), *Foundations of Augmented Cognition. Advancing Human Performance and Decision-Making through Adaptive Systems* (Vol. 8534, pp. 3–13). Springer International Publishing. https://doi.org/10.1007/978-3-319-07527-3_1

Bradley, M. M. (2009). Natural selective attention: Orienting and emotion. *Psychophysiology*, *46*(1), 1–11. https://doi.org/10.1111/j.1469-8986.2008.00702.x

Bradley, M. M., & Lang, P. J. (2000). Affective reactions to acoustic stimuli. *Psychophysiology*, *37*(2), 204–215. https://doi.org/10.1111/1469-8986.3720204

Brattico, E. (2021). The Empirical Aesthetics of Music. In M. Nadal & O. Vartanian (Eds.), *The Oxford Handbook of Empirical Aesthetics* (1st ed., pp. 573–604). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198824350.013.26

Brattico, E., Alluri, V., Bogert, B., Jacobsen, T., Vartiainen, N., Nieminen, S., & Tervaniemi, M. (2011). A Functional MRI Study of Happy and Sad Emotions in Music with and without Lyrics. *Frontiers in Psychology*, *2*. https://doi.org/10.3389/fpsyg.2011.00308

Brattico, E., Bogert, B., & Jacobsen, T. (2013). Toward a Neural Chronometry for the Aesthetic Experience of Music. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00206

Brattico, E., Jacobsen, T., De Baene, W., Glerean, E., & Tervaniemi, M. (2010). Cognitive vs. Affective listening modes and judgments of music – An ERP study. *Biological Psychology*, *85*(3), 393–409. https://doi.org/10.1016/j.biopsycho.2010.08.014

Brattico, E., & Pearce, M. (2013). The neuroaesthetics of music. *Psychology of Aesthetics, Creativity, and the Arts*, *7*(1), 48–61. https://doi.org/10.1037/a0031624

Broughton, M., & Stevens, C. (2009). Music, movement and marimba: An investigation of the role of movement and gesture in communicating musical expression to an audience. *Psychology of Music*, *37*(2), 137–153. https://doi.org/10.1177/0305735608094511

Brown, P., Rothwell, J. C., Thompson, P. D., Britton, T. C., Day, B. L., & Marsden, C. D. (1991). New obervations on the normal auditory startle reflex in man. *Brain*, *114*(4), 1891–1902. https://doi.org/10.1093/brain/114.4.1891

Brown, R., Kemp, U., & Macefield, V. (2013). Increases in muscle sympathetic nerve activity, heart rate, respiration, and skin blood flow during passive viewing of exercise. *Frontiers in Neuroscience*, *7*. https://doi.org/10.3389/fnins.2013.00102

Brown, S. C., & Knox, D. (2017). Why go to pop concerts? The motivations behind live music attendance. *Musicae Scientiae*, *21*(3), 233–249. https://doi.org/10.1177/1029864916650719

Burger, B., Saarikallio, S., Luck, G., Thompson, M. R., & Toiviainen, P. (2013). Relationships Between Perceived Emotions in Music and Music-induced Movement. *Music Perception*, *30*(5), 517–533. https://doi.org/10.1525/mp.2013.30.5.517

Burland, K., & Pitts, S. (Eds.). (2014). *Coughing and clapping: Investigating audience experience*. Ashgate Publishing Group.

Burunat, I., Alluri, V., Toiviainen, P., Numminen, J., & Brattico, E. (2014). Dynamics of brain activity underlying working memory for music in a naturalistic condition. *Cortex*, *57*, 254–269. https://doi.org/10.1016/j.cortex.2014.04.012

Burunat, I., Toiviainen, P., Alluri, V., Bogert, B., Ristaniemi, T., Sams, M., & Brattico, E. (2016). The reliability of continuous brain responses during naturalistic listening to music. *NeuroImage*, *124*, 224–231. https://doi.org/10.1016/j.neuroimage.2015.09.005

Cacioppo, J. T., Berntson, G. G., Larsen, J. T., Poehlmann, K. M., & Ito, T. A. (2000). The Psychophysiology of Emotion. In *The Handbook of Emotion* (2nd ed., pp. 173–191). Guildford Press.

Camurri, A., Mazzarino, B., Ricchetti, M., Timmers, R., & Volpe, G. (2004). Multimodal Analysis of Expressive Gesture in Music and Dance Performances. In A. Camurri & G. Volpe (Eds.), *Gesture-Based Communication in Human-Computer Interaction* (Vol. 2915, pp. 20–39). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-24598-8_3

Cannam, C., Landone, C., & Sandler, M. (2010). Sonic visualiser: An open source application for viewing, analysing, and annotating music audio files. *Proceedings of the 18th ACM International Conference on Multimedia*, 1467–1468.

Caplin, W. E. (1998). *Classical form: A theory of formal functions for the instrumental music of Haydn, Mozart, and Beethoven.* Oxford University Press.

Cardona, G., Ferreri, L., Lorenzo-Seva, U., Russo, F. A., & Rodriguez-Fornells, A. (2022). The forgotten role of absorption in music reward. *Annals of the New York Academy of Sciences*, *1514*(1), 142–154. https://doi.org/10.1111/nyas.14790

Chabin, T., Gabriel, D., Chansophonkul, T., Michelant, L., Joucla, C., Haffen, E., Moulin, T., Comte, A., & Pazart, L. (2020). Cortical Patterns of Pleasurable Musical Chills Revealed by High-Density EEG. *Frontiers in Neuroscience*, *14*, 565815. https://doi.org/10.3389/fnins.2020.565815

Chabin, T., Gabriel, D., Comte, A., Haffen, E., Moulin, T., & Pazart, L. (2022). Interbrain emotional connection during music performances is driven by physical proximity and individual traits. *Annals of the New York Academy of Sciences*, *1508*(1), 178–195. https://doi.org/10.1111/nyas.14711

Chaffin, R., Lemieux, A. F., & Chen, C. (2007). 'It is Different Each Time I Play': Variability in Highly Prepared Musical Performance. *Music Perception*, *24*(5), 455–472. https://doi.org/10.1525/mp.2007.24.5.455

Chalas, N., Daube, C., Kluger, D. S., Abbasi, O., Nitsch, R., & Gross, J. (2023). Speech onsets and sustained speech contribute differentially to delta and theta speech tracking in auditory cortex. *Cerebral Cortex*, bhac502. https://doi.org/10.1093/cercor/bhac502

Chamorro-Premuzic, T., & Furnham, A. (2007). Personality and music: Can traits explain how people use music in everyday life? *British Journal of Psychology*, *98*(2), 175–185. https://doi.org/10.1348/000712606X111177

Chan, L. P., Livingstone, S. R., & Russo, F. A. (2013). Facial Mimicry in Response to Song. *Music Perception*, *30*(4), 361–367. https://doi.org/10.1525/mp.2013.30.4.361

Chapados, C., & Levitin, D. J. (2008). Cross-modal interactions in the experience of musical performances: Physiological correlates. *Cognition*, *108*(3), 639–651. https://doi.org/10.1016/j.cognition.2008.05.008

Cheung, V. K. M., Harrison, P. M. C., Meyer, L., Pearce, M. T., Haynes, J.-D., & Koelsch, S. (2019). Uncertainty and Surprise Jointly Predict Musical Pleasure and Amygdala, Hippocampus, and Auditory Cortex Activity. *Current Biology*, *29*(23), 4084-4092.e4. https://doi.org/10.1016/j.cub.2019.09.067

Chin, T., & Rickard, N. S. (2012). The Music USE (MUSE) Questionnaire: An Instrument to Measure Engagement in Music. *Music Perception*, *29*(4), 429–446. https://doi.org/10.1525/mp.2012.29.4.429

Christensen, J. F., Azevedo, R. T., & Tsakiris, M. (2021). Emotion matters: Different psychophysiological responses to expressive and non-expressive full-body movements. *Acta Psychologica*, *212*, 103215. https://doi.org/10.1016/j.actpsy.2020.103215

Chuen, L., Sears, D., & McAdams, S. (2016). Psychophysiological responses to auditory change: Psychophysiological responses to auditory change. *Psychophysiology*, *53*(6), 891–904. https://doi.org/10.1111/psyp.12633

Clarke, E. F., & Krumhansl, C. L. (1990). Perceiving Musical Time. *Music Perception: An Interdisciplinary Journal*, *7*(3), 213–251. https://doi.org/10.2307/40285462

Cohen. (1988). Set Correlation and Contingency Tables. *Applied Psychological Measurement*, *12*(4), 425–434. https://doi.org/10.1177/014662168801200410

Cohen. (2014). *Analyzing neural time series data: Theory and practice*. The MIT Press.

Cohen, A. (2009). Music in performance arts: Film, theatre and dance. In S. Hallam, I. Cross, & M. Thaut (Eds.), *The Oxford handbook of music psychology* (1st ed.). Oxford University Press.

Cohen, & Parra, L. C. (2016). Memorable Audiovisual Narratives Synchronize Sensory and Supramodal Neural Responses. *Eneuro*, *3*(6), ENEURO.0203-16.2016. https://doi.org/10.1523/ENEURO.0203-16.2016

Connolly, S. (2022, November 28). An open letter from Dame Sarah Connolly and colleagues to Arts Council England. *The Arts Desk*. https://theartsdesk.com/opera/open-letter-dame-sarah-connolly-and-colleagues-arts-council-england

Cook, N. (1987). The Perception of Large-Scale Tonal Closure. *Music Perception*, *5*(2), 197–205. https://doi.org/10.2307/40285392

Coutinho, E., & Cangelosi, A. (2011). Musical emotions: Predicting second-by-second subjective feelings of emotion from low-level psychoacoustic features and physiological measurements. *Emotion*, *11*(4), 921–937. https://doi.org/10.1037/a0024700

Coutinho, E., & Scherer, K. R. (2017a). Introducing the GEneva Music-Induced Affect Checklist (GEMIAC). *Music Perception*, *34*(4), 371–386. https://doi.org/10.1525/mp.2017.34.4.371

Coutinho, E., & Scherer, K. R. (2017b). The effect of context and audio-visual modality on emotions elicited by a musical performance. *Psychology of Music*, *45*(4), 550–569. https://doi.org/10.1177/0305735616670496

Craig, D. G. (2005). An Exploratory Study of Physiological Changes during "Chills" Induced by Music. *Musicae Scientiae*, *9*(2), 273–287. https://doi.org/10.1177/102986490500900207

Criscuolo, A., Schwartze, M., & Kotz, S. A. (2022). Cognition through the lens of a body–brain dynamic system. *Trends in Neurosciences*, S0166223622001229. https://doi.org/10.1016/j.tins.2022.06.004

Crocker, R. L. (1963). Pythagorean Mathematics and Music. *The Journal of Aesthetics and Art Criticism*, *22*(2), 189. https://doi.org/10.2307/427754

Cross, E. (2011). The impact of aesthetic evaluation and physical ability on dance perception. *Frontiers in Human Neuroscience*, *5*. https://doi.org/10.3389/fnhum.2011.00102

Cross, E. (2015). Beautiful embodiment: The shaping of aesthetic preference by personal experience. In *Art, Aesthetics and the Brain.* (pp. 189-208.). Oxford University Press, Oxford.

Cupchik, G. C. (2019). COMMENT Reflections on Relations Between Affective Responses and Operatic Intentions: Comment on Scherer, Trznadel, Fantini, and Coutinho (2019). *Psychology of Aesthetics, Creativity, and the Arts*, *13*(3).

Czepiel, A. (2022). *R demo: Basic cardiorespiratory analysis (Nijmegen, 2022-10)*. https://osf.io/a7yfn/

Czepiel, A., Fink, L. K., Seibert, C., Scharinger, M., & Kotz, S. A. (2023). *Aesthetic and physiological effects of naturalistic multimodal music listening* [Preprint]. biorxiv. https://doi.org/10.1101/2022.07.02.498390

Czepiel, A., Fink, L. K., Fink, L. T., Wald-Fuhrmann, M., Tröndle, M., & Merrill, J. (2021). Synchrony in the periphery: Inter-subject correlation of physiological responses during live music concerts. *Scientific Reports*, *11*(1), 22457. https://doi.org/10.1038/s41598-021-00492-3

Czerny, C. (1930). *School of practical composition (J.Bishop, Trans, original work published 1848).*

Dahl, S., & Friberg, A. (2007). Visual Perception of Expressiveness in Musicians' Body Movements. *Music Perception*, *24*(5), 433–454. https://doi.org/10.1525/mp.2007.24.5.433

Dauer, T., Nguyen, D. T., Gang, N., Dmochowski, J. P., Berger, J., & Kaneshiro, B. (2021). Inter-subject Correlation While Listening to Minimalist Music: A Study of Electrophysiological and Behavioral Responses to Steve Reich's Piano Phase. *Frontiers in Neuroscience*, *15*. https://www.frontiersin.org/articles/10.3389/fnins.2021.702067

Davidson, J. W. (1993). Visual Perception of Performance Manner in the Movements of Solo Musicians. *Psychology of Music*, *21*(2), 103–113. https://doi.org/10.1177/030573569302100201

Davidson, J. W. (2012). Bodily movement and facial actions in expressive musical performance by solo and duo instrumentalists: Two distinctive case studies. *Psychology of Music*, *40*(5), 595–633. https://doi.org/10.1177/0305735612449896

Davis, R. C. (1934). Modification of the galvanic reflex by daily repetition of a stimulus. *Journal of Experimental Psychology*, *17*(4), 504–535. https://doi.org/10.1037/h0074305

Dawson, M. E., Schell, A. M., & Filion, D. L. (2016). The Electrodermal System. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of Psychophysiology* (4th ed., pp. 217–243). Cambridge University Press. https://doi.org/10.1017/9781107415782.010

de Jong, M. A., van Mourik, K. R., & Schellekens, H. M. C. (1973). A Physiological Approach to Aesthetic Preference. *Psychotherapy and Psychosomatics*, *22*(1), 46–51. https://doi.org/10.1159/000286468

de Morree, H. M., & Marcora, S. M. (2010). The face of effort: Frowning muscle activity reflects effort during a physical task. *Biological Psychology*, *85*(3), 377–382. https://doi.org/10.1016/j.biopsycho.2010.08.009

Dearn, L. K., & Pitts, S. E. (2017). (Un)popular music and young audiences: Exploring the classical chamber music concert from the perspective of young adult listeners. *Journal of Popular Music Education*, *1*(1), 43–62. https://doi.org/10.1386/jpme.1.1.43_1

Dearn, L. K., & Price, S. M. (2016). Sharing Music: Social and Communal Aspects of Concert-Going. *Networking Knowledge: Journal of the MeCCSA Postgraduate Network*, *9*(2). https://doi.org/10.31165/nk.2016.92.428

Deliege, I. (1987). Grouping Conditions in Listening to Music: An Approach to Lerdahl & Jackendoff's Grouping Preference Rules. *Music Perception: An Interdisciplinary Journal*, *4*(4), 325–359. https://doi.org/10.2307/40285378

Dellacherie, D., Roy, M., Hugueville, L., Peretz, I., & Samson, S. (2011). The effect of musical experience on emotional self-reports and psychophysiological responses to dissonance: Psychophysiology of musical emotion. *Psychophysiology*, *48*(3), 337–349. https://doi.org/10.1111/j.1469-8986.2010.01075.x

Deutscher Bühnenverein. (2022). *Events and attendance at publicly financed music theatres.* https://miz.org/en/statistics/concerts-and-attendance-of-publicly-financed-orchestras

Di Bernardi Luft, C., & Bhattacharya, J. (2015). Aroused with heart: Modulation of heartbeat evoked potential by arousal induction and its oscillatory correlates. *Scientific Reports*, *5*(1), 15717. https://doi.org/10.1038/srep15717

Di Liberto, G. M., Pelofi, C., Bianco, R., Patel, P., Mehta, A. D., Herrero, J. L., de Cheveigné, A., Shamma, S., & Mesgarani, N. (2020). Cortical encoding of melodic expectations in human temporal cortex. *ELife*, *9*, e51784. https://doi.org/10.7554/eLife.51784

Dikker, S., Michalareas, G., Oostrik, M., Serafimaki, A., Kahraman, H. M., Struiksma, M. E., & Poeppel, D. (2021). Crowdsourcing neuroscience: Inter-brain coupling during face-to-face interactions outside the laboratory. *NeuroImage*, *227*, 117436. https://doi.org/10.1016/j.neuroimage.2020.117436

Dikker, S., Wan, L., Davidesco, I., Kaggen, L., Oostrik, M., McClintock, J., Rowland, J., Michalareas, G., Van Bavel, J. J., Ding, M., & Poeppel, D. (2017). Brain-to-Brain Synchrony Tracks Real-World Dynamic Group Interactions in the Classroom. *Current Biology*, *27*(9), 1375–1380. https://doi.org/10.1016/j.cub.2017.04.002

Dillman Carpentier, F. R., & Potter, R. F. (2007). Effects of Music on Physiological Arousal: Explorations into Tempo and Genre. *Media Psychology*, *10*(3), 339–363. https://doi.org/10.1080/15213260701533045

Dimberg, U. (1982). Facial Reactions to Facial Expressions. *Psychophysiology*, *19*(6), 643–647. https://doi.org/10.1111/j.1469-8986.1982.tb02516.x

Dimberg, U. (1990). Facial electromyographic reactions and autonomic activity to auditory stimuli. *Biological Psychology*, *31*(2), 137–147. https://doi.org/10.1016/0301-0511(90)90013-M

Dimberg, U., Thunberg, M., & Elmehed, K. (2000). Unconscious Facial Reactions to Emotional Facial Expressions. *Psychological Science*, *11*(1), 86–89. https://doi.org/10.1111/1467-9280.00221

Ding, N., Pan, X., Luo, C., Su, N., Zhang, W., & Zhang, J. (2018). Attention Is Required for Knowledge-Based Sequential Grouping: Insights from the Integration of Syllables into Words. *The Journal of Neuroscience*, *38*(5), 1178–1188. https://doi.org/10.1523/JNEUROSCI.2606-17.2017

Ding, N., Patel, A. D., Chen, L., Butler, H., Luo, C., & Poeppel, D. (2017). Temporal modulations in speech and music. *Neuroscience & Biobehavioral Reviews*, *81*, 181–187. https://doi.org/10.1016/j.neubiorev.2017.02.011

Dmochowski, J. P., Sajda, P., Dias, J., & Parra, L. C. (2012). Correlated Components of Ongoing EEG Point to Emotionally Laden Attention – A Possible Marker of Engagement? *Frontiers in Human Neuroscience*, *6*. https://doi.org/10.3389/fnhum.2012.00112

Doelling, K. B., & Poeppel, D. (2015). Cortical entrainment to music and its modulation by expertise. *Proceedings of the National Academy of Sciences*, *112*(45), E6233–E6242. https://doi.org/10.1073/pnas.1508431112

Dotov, D., & Trainor, L. J. (2021). Cross-frequency coupling explains the preference for simple ratios in rhythmic behaviour and the relative stability across non-synchronous patterns. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *376*(1835), 20200333. https://doi.org/10.1098/rstb.2020.0333

Drabkin, W. (2001). Sequence (ii). In *Oxford Music Online*. Oxford University Press. https://doi.org/10.1093/gmo/9781561592630.article.25439

Eerola, T., Vuoskoski, J. K., & Kautiainen, H. (2016). Being Moved by Unfamiliar Sad Music Is Associated with High Empathy. *Frontiers in Psychology*, *7*, 1176. https://doi.org/10.3389/fpsyg.2016.01176

Eerola, T., Vuoskoski, J. K., Peltola, H.-R., Putkinen, V., & Schäfer, K. (2018). An integrative review of the enjoyment of sadness associated with music. *Physics of Life Reviews*, *25*, 100–121. https://doi.org/10.1016/j.plrev.2017.11.016

Egermann, H., Fernando, N., Chuen, L., & McAdams, S. (2015). Music induces universal emotion-related psychophysiological responses: Comparing Canadian listeners to Congolese Pygmies. *Frontiers in Psychology*, *5*. https://www.frontiersin.org/article/10.3389/fpsyg.2014.01341

Egermann, H., & McAdams, S. (2013). Empathy and Emotional Contagion as a Link Between Recognized and Felt Emotions in Music Listening. *Music Perception*, *31*(2), 139–156. https://doi.org/10.1525/mp.2013.31.2.139

Egermann, H., Pearce, M. T., Wiggins, G. A., & McAdams, S. (2013). Probabilistic models of expectation violation predict psychophysiological emotional responses to live concert music. *Cognitive, Affective, & Behavioral Neuroscience*, *13*(3), 533–553. https://doi.org/10.3758/s13415-013-0161-y

Egermann, H., & Reuben, F. (2020). "Beauty Is How You Feel Inside": Aesthetic Judgments Are Related to Emotional Responses to Contemporary Music. *Frontiers in Psychology*, *11*. https://www.frontiersin.org/article/10.3389/fpsyg.2020.510029

Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion*, *6*(3–4), 169–200. https://doi.org/10.1080/02699939208411068

Ekman, P., & Friesen, W. V. (1978). *Facial Action Coding System (FACS)* [Data set]. American Psychological Association. https://doi.org/10.1037/t27734-000

Ekman, P., Levenson, R. W., & Friesen, W. V. (1983). Autonomic Nervous System Activity Distinguishes Among Emotions. *Science*, *221*(4616), 1208–1210. https://doi.org/10.1126/science.6612338

Ellis, R. J., & Thayer, J. F. (2010). Music and Autonomic Nervous System (Dys)Function. *Music Perception*, *27*(4), 317–326. https://doi.org/10.1525/mp.2010.27.4.317

Etzel, J. A., Johnsen, E. L., Dickerson, J., Tranel, D., & Adolphs, R. (2006). Cardiovascular and respiratory responses during musical mood induction. *International Journal of Psychophysiology*, *61*(1), 57–69. https://doi.org/10.1016/j.ijpsycho.2005.10.025

Evans, P., & Schubert, E. (2008). Relationships between expressed and felt emotions in music. *Musicae Scientiae*, *12*(1), 75–99. https://doi.org/10.1177/102986490801200105

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, *4*(3), 272–299. https://doi.org/10.1037/1082-989X.4.3.272

Fechner, G. T. (1876). *Vorschule der Ästhetik [Preliminaries of Aesthetics]*. Breitkopf und Härtel.

Fernández-Rubio, G., Brattico, E., Kotz, S. A., Kringelbach, M. L., Vuust, P., & Bonetti, L. (2022). Magnetoencephalography recordings reveal the spatiotemporal dynamics of recognition memory for complex versus simple auditory sequences. *Communications Biology*, *5*(1), 1272. https://doi.org/10.1038/s42003-022-04217-8

Finisguerra, A., Ticini, L. F., Kirsch, L. P., Cross, E. S., Kotz, S. A., & Urgesi, C. (2021). Dissociating embodiment and emotional reactivity in motor responses to artworks. *Cognition*, *212*, 104663. https://doi.org/10.1016/j.cognition.2021.104663

Fink, L. K., Hurley, B. K., Geng, J. J., & Janata, P. (2018). A linear oscillator model predicts dynamic temporal attention and pupillary entrainment to rhythmic patterns. *Journal of Eye Movement Research*, *11*(2), 10.16910/jemr.11.2.12. https://doi.org/10.16910/jemr.11.2.12

Fink, L. K., Simola, J., Tavano, A., Lange, E., Wallot, S., & Laeng, B. (2023). From pre-processing to advanced dynamic modeling of pupil data. *Behavior Research Methods. Preprint on PsyArXiv: Https://Psyarxiv.Com/Wqvue.*

Fink, L. K., Warrenburg, L. A., Howlin, C., Randall, W. M., Hansen, N. Chr., & Wald-Fuhrmann, M. (2021). Viral tunes: Changes in musical behaviours and interest in coronamusic predict socio-emotional coping during COVID-19 lockdown. *Humanities and Social Sciences Communications*, *8*(1), 180. https://doi.org/10.1057/s41599-021-00858-y

Fowles, D. C., Christie, M. J., Edelberg, R., Grings, W. W., Lykken, D. T., & Venables, P. H. (1981). Publication Recommendations for Electrodermal Measurements. *Psychophysiology*, *18*(3), 232–239. https://doi.org/10.1111/j.1469-8986.1981.tb03024.x

Fox, J., & Weisberg, S. (2019). *An R Companion to Applied Regression* (Third). https://socialsciences.mcmaster.ca/jfox/Books/Companion/

Fredrickson, B. L. (2000). Extracting meaning from past affective experiences: The importance of peaks, ends, and specific emotions. *Cognition & Emotion*, *14*(4), 577–606. https://doi.org/10.1080/026999300402808

Freedberg, D., & Gallese, V. (2007). Motion, emotion and empathy in esthetic experience. *Trends in Cognitive Sciences*, *11*(5), 197–203. https://doi.org/10.1016/j.tics.2007.02.003

Frith, C. D., & Allen, H. A. (1983). The skin conductance orienting response as an index of attention. *Biological Psychology*, *17*(1), 27–39. https://doi.org/10.1016/0301-0511(83)90064-9

Frobenius, W., Cooke, P., Bithell, C., & Zemtsovsky, I. (2001). *Polyphony* (Vol. 1). Oxford University Press. https://doi.org/10.1093/gmo/9781561592630.article.42927

Fuentes-Sánchez, N., Pastor, R., Eerola, T., Escrig, M. A., & Pastor, M. C. (2022). Musical preference but not familiarity influences subjective ratings and psychophysiological correlates of music-induced emotions. *Personality and Individual Differences*, *198*, 111828. https://doi.org/10.1016/j.paid.2022.111828

Gabrielsson, A. (2002). Emotion Perceived and Emotion Felt: Same or Different? *Musicae Scientiae*, *5*(1 suppl), 123–147. https://doi.org/10.1177/10298649020050S105

Gabrielsson, A. (2012). *The relationship between musical structure and perceived expression* (S. Hallam, I. Cross, & M. Thaut, Eds.; Vol. 1). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199298457.013.0013

Gabrielsson, A., & Juslin, P. N. (1996). Emotional Expression in Music Performance: Between the Performer's Intention and the Listener's Experience. *Psychology of Music*, *24*(1), 68–91. https://doi.org/10.1177/0305735696241007

Gabrielsson, A., & Wik, S. L. (2003). Strong Experiences Related to Music: A Descriptive System. *Musicae Scientiae*, *7*(2), 157–217. https://doi.org/10.1177/102986490300700201

Gallese, V., & Freedberg, D. (2007). Mirror and canonical neurons are crucial elements in esthetic response. *Trends in Cognitive Sciences*, *11*(10), 411. https://doi.org/10.1016/j.tics.2007.07.006

Gernot, G., Pelowski, M., & Leder, H. (2018). Empathy, Einfühlung, and aesthetic experience: The effect of emotion contagion on appreciation of representational and abstract art using fEMG and SCR. *Cognitive Processing*, *19*(2), 147–165. https://doi.org/10.1007/s10339-017-0800-2

Gilmore, S. (1993). Tradition and novelty in concert programming: Bringing the artist back into cultural analysis. *Sociological Forum*, *8*(2), 221–242. https://doi.org/10.1007/BF01115491

Gingras, B., Marin, M. M., & Fitch, W. T. (2014). Beyond Intensity: Spectral Features Effectively Predict Music-Induced Subjective Arousal. *Quarterly Journal of Experimental Psychology*, *67*(7), 1428–1446. https://doi.org/10.1080/17470218.2013.863954

Giordano, B. L., Whiting, C., Kriegeskorte, N., Kotz, S. A., Gross, J., & Belin, P. (2021). The representational dynamics of perceived voice emotions evolve from categories to dimensions. *Nature Human Behaviour*, *5*(9), 1203–1213. https://doi.org/10.1038/s41562-021-01073-0

Goffman, E. (1974). *Frame analysis: An essay on the organization of experience*. Harvard University Press.

Goldstein, A. (1980). Thrills in response to music and other stimuli. *Physiological Psychology*, *8*(1), 126–129. https://doi.org/10.3758/BF03326460

Golland, Y., Arzouan, Y., & Levit-Binnun, N. (2015). The Mere Co-Presence: Synchronization of Autonomic Signals and Emotional Responses across Co-Present Individuals Not Engaged in Direct Interaction. *PLOS ONE*, *10*(5), e0125804. https://doi.org/10.1371/journal.pone.0125804

Gomez, P., & Danuser, B. (2004). Affective and physiological responses to environmental noises and music. *International Journal of Psychophysiology*, *53*(2), 91–103. https://doi.org/10.1016/j.ijpsycho.2004.02.002

Gomez, P., & Danuser, B. (2007). Relationships between musical structure and psychophysiological measures of emotion. *Emotion*, *7*, 377–387. https://doi.org/10.1037/1528-3542.7.2.377

Gonzalez-Sanchez, V. E., Zelechowska, A., & Jensenius, A. R. (2018). Correspondences Between Music and Involuntary Human Micromotion During Standstill. *Frontiers in Psychology*, *9*, 1382. https://doi.org/10.3389/fpsyg.2018.01382

Goodchild, M., Wild, J., & McAdams, S. (2019). Exploring emotional responses to orchestral gestures. *Musicae Scientiae*, *23*(1), 25–49. https://doi.org/10.1177/1029864917704033

Gordan, R., Gwathmey, J. K., & Xie, L.-H. (2015). Autonomic and endocrine control of cardiovascular function. *World Journal of Cardiology*, *7*(4), 204. https://doi.org/10.4330/wjc.v7.i4.204

Gorzelańczyk, E. J., Podlipniak, P., Walecki, P., Karpiński, M., & Tarnowska, E. (2017). Pitch Syntax Violations Are Linked to Greater Skin Conductance Changes, Relative to Timbral Violations – The Predictive Role of the Reward System in Perspective of Cortico–subcortical Loops. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.00586

Gotham, M. (2014). Coherence in Concert Programming: A View from the U.K. *International Review of the Aesthetics and Sociology of Music*, *45*(2), 293–309.

Gotlieb, H., & Konečni, V. J. (1985). The Effects of Instrumentation, Playing Style, and Structure in the Goldberg Variations by Johann Sebastian Bach. *Music Perception*, *3*(1), 87–101. https://doi.org/10.2307/40285323

Grace, A. (2020). *Dutch concert revenue up 25% in 2019*. https://www.iq-mag.net/2020/01/concert-revenue-up-25-in-2019-netherlands/

Graham, F. K., & Clifton, R. K. (1966). Heart-rate change as a component of the orienting response. *Psychological Bulletin*, *65*(5), 305–320. https://doi.org/10.1037/h0023258

Grahn, J. A., & Brett, M. (2007). Rhythm and Beat Perception in Motor Areas of the Brain. *Journal of Cognitive Neuroscience*, *19*(5), 893–906. https://doi.org/10.1162/jocn.2007.19.5.893

Greb, F., Schlotz, W., & Steffens, J. (2018). Personal and situational influences on the functions of music listening. *Psychology of Music*, *46*(6), 763–794. https://doi.org/10.1177/0305735617724883

Grewe, O., Kopiez, R., & Altenmüüller, E. (2009). The Chill Parameter: Goose Bumps and Shivers as Promising Measures in Emotion Research. *Music Perception*, *27*(1), 61–74. https://doi.org/10.1525/mp.2009.27.1.61

Grewe, O., Nagel, F., Kopiez, R., & Altenmüller, E. (2007). Listening To Music As A Re-Creative Process: Physiological, Psychological, And Psychoacoustical Correlates Of Chills And Strong Emotions. *Music Perception: An Interdisciplinary Journal*, *24*(3), 297–314. https://doi.org/10.1525/mp.2007.24.3.297

Grey, J. M., & Gordon, J. W. (1978). Perceptual effects of spectral modifications on musical timbres. *The Journal of the Acoustical Society of America*, *63*(5), 1493–1500. https://doi.org/10.1121/1.381843

Griffiths, N. K. (2010). 'Posh music should equal posh dress': An investigation into the concert dress and physical appearance of female soloists. *Psychology of Music*, *38*(2), 159–177. https://doi.org/10.1177/0305735608100372

Griffiths, N. K. (2011). The fabric of performance: Values and social practices of classical music expressed through concert dress choice. *Music Performance Research*, *4*, 30–48.

Griffiths, N. K., & Reay, J. L. (2018). The Relative Importance of Aural and Visual Information in the Evaluation of Western Canon Music Performance by Musicians and Nonmusicians. *Music Perception*, *35*(3), 364–375. https://doi.org/10.1525/mp.2018.35.3.364

Grove Music Online. (2001). *Appoggiatura*. Oxford University Press. https://doi.org/10.1093/gmo/9781561592630.article.01118

Grund, M., Al, E., Pabst, M., Dabbagh, A., Stephani, T., Nierhaus, T., Gaebler, M., & Villringer, A. (2022). Respiration, Heartbeat, and Conscious Tactile Perception. *The Journal of Neuroscience*, *42*(4), 643–656. https://doi.org/10.1523/JNEUROSCI.0592-21.2021

Gugnowska, K., Novembre, G., Kohler, N., Villringer, A., Keller, P. E., & Sammler, D. (2022). Endogenous sources of interbrain synchrony in duetting pianists. *Cerebral Cortex*, *32*(18), 4110–4127. https://doi.org/10.1093/cercor/bhab469

Guhn, M., Hamm, A., & Zentner, M. (2007). Physiological and Musico-Acoustic Correlates of the Chill Response. *Music Perception*, *24*(5), 473–484. https://doi.org/10.1525/mp.2007.24.5.473

Gupta, U., & Gupta, B. S. (2015). Psychophysiological reactions to music in male coronary patients and healthy controls. *Psychology of Music*, *43*(5), 736–755. https://doi.org/10.1177/0305735614536754

Haas, F., Distenfeld, S., & Axen, K. (1986). Effects of perceived musical rhythm on respiratory pattern. *Journal of Applied Physiology*, *61*(3), 1185–1191. https://doi.org/10.1152/jappl.1986.61.3.1185

Hall, S. E., Schubert, E., & Wilson, S. J. (2016). The Role of Trait and State Absorption in the Enjoyment of Music. *PLOS ONE*, *11*(11), e0164029. https://doi.org/10.1371/journal.pone.0164029

Halpern, A., Chan, C., Müllensiefen, D., & Sloboda, J. (2017). Audience Reactions to Repeating A Piece on A Concert Programme. *Participations*, *14*(2), 135–152.

Hamann, S. (2012). Mapping discrete and dimensional emotions onto the brain: Controversies and consensus. *Trends in Cognitive Sciences*, *16*(9), 458–466. https://doi.org/10.1016/j.tics.2012.07.006

Harding, E. E., Sammler, D., Henry, M. J., Large, E. W., & Kotz, S. A. (2019). Cortical tracking of rhythm in music and speech. *NeuroImage*, *185*, 96–101. https://doi.org/10.1016/j.neuroimage.2018.10.037

Hargreaves, D. J., & North, A. C. (2010). Experimental aesthetics and liking for music. In *Handbook of Music and Emotion: Theory, Research, Applications Handbook of Music and Emotion: Theory, Research, Applications* (pp. 515–546). Oxford University Press, Oxford. https://academic.oup.com/book/38621/chapter/335188577

Harrison, Bianco, R., Chait, M., & Pearce, M. T. (2020). PPM-Decay: A computational model of auditory prediction with memory decay. *PLOS Computational Biology*, *16*(11), e1008304. https://doi.org/10.1371/journal.pcbi.1008304

Harrison, & Loui, P. (2014). Thrills, chills, frissons, and skin orgasms: Toward an integrative model of transcendent psychophysiological experiences in music. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00790

Hartmann, M., Mavrolampados, A., Allingham, E., Carlson, E., Burger, B., & Toiviainen, P. (2019). Kinematics of perceived dyadic coordination in dance. *Scientific Reports*, *9*(1), 15594. https://doi.org/10.1038/s41598-019-52097-6

Hasson, U., Malach, R., & Heeger, D. J. (2010). Reliability of cortical activity during natural stimulation. *Trends in Cognitive Sciences*, *14*(1), 40–48. https://doi.org/10.1016/j.tics.2009.10.011

Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., & Malach, R. (2004). Intersubject Synchronization of Cortical Activity During Natural Vision. *Science*, *303*(5664), 1634–1640. https://doi.org/10.1126/science.1089506

Hausfeld, L., Disbergen, N. R., Valente, G., Zatorre, R. J., & Formisano, E. (2021). Modulating Cortical Instrument Representations During Auditory Stream Segregation and Integration With Polyphonic Music. *Frontiers in Neuroscience*, *15*, 635937. https://doi.org/10.3389/fnins.2021.635937

Hausfeld, L., Riecke, L., Valente, G., & Formisano, E. (2018). Cortical tracking of multiple streams outside the focus of attention in naturalistic auditory scenes. *NeuroImage*, *181*, 617–626. https://doi.org/10.1016/j.neuroimage.2018.07.052

Heister, H.-W. (1983). *Das Konzert: Theorie einer Kulturform: Vol. Band 1*. Heinrichshofen.

Heister, H.-W. (1984). Geldloses Geschenk und archaisches Zeremoniell. Der Konzert-Beifall als Honorar- und Aktivitätsform. *International Review of the Aesthetics and Sociology of Music*, *15*(2), 91. https://doi.org/10.2307/836859

Helmholtz, H. von, & Ellis, A. J. (1875). *On the sensations of tone as a physiological basis for the theory of music*. Longmans, Green and Company.

Henry, M. J., & Obleser, J. (2012). Frequency modulation entrains slow neural oscillations and optimizes human listening behavior. *Proceedings of the National Academy of Sciences*, *109*(49), 20095–20100. https://doi.org/10.1073/pnas.1213390109

Hepokoski, J., & Darcy, W. (2006). *Elements of Sonata Theory*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195146400.001.0001

Herbec, A., Kauppi, J.-P., Jola, C., Tohka, J., & Pollick, F. E. (2015). Differences in fMRI intersubject correlation while viewing unedited and edited videos of dance performance. *Cortex*, *71*, 341–348. https://doi.org/10.1016/j.cortex.2015.06.026

Herbert, R. (2016). *Everyday music listening: Absorption, dissociation and trancing*. Routledge Taylor & Francis group.

Herrmann, B., & Johnsrude, I. S. (2020). Absorption and Enjoyment During Listening to Acoustically Masked Stories. *Trends in Hearing*, *24*, 233121652096785. https://doi.org/10.1177/2331216520967850

Hevner, K. (1936). Experimental Studies of the Elements of Expression in Music. *The American Journal of Psychology*, *48*(2), 246. https://doi.org/10.2307/1415746

Hewett, I. (2017). Europe. In *The other classical musics: Fifteen great traditions. Yearbook for traditional music* (Vol. 49, pp. 216–245). Boydell and Brewer.

Hodges, D. (2009). Bodily responses to music. In *The Oxford Handbook of Music Psychology* (pp. 121–130).

Hodges, D. (2010). Psychophysiological Measures. In *Handbook of Music and Emotion: Theory, Research, Applications* (pp. 278–311). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199230143.001.0001

Høffding, S. (2018). *A Phenomenology of Musical Absorption*. Springer International Publishing. https://doi.org/10.1007/978-3-030-00659-4

Hogue, J. D., Crimmins, A. M., & Kahn, J. H. (2016). "So sad and slow, so why can't I turn off the radio": The effects of gender, depression, and absorption on liking music that induces sadness and music that induces happiness. *Psychology of Music*, *44*(4), 816–829. https://doi.org/10.1177/0305735615594489

Holroyd, C. B. (2022). Interbrain synchrony: On wavy ground. *Trends in Neurosciences*, *45*(5), 346–357. https://doi.org/10.1016/j.tins.2022.02.002

Holt, F. (2010). The economy of live music in the digital age. *European Journal of Cultural Studies*, *13*(2), 243–261. https://doi.org/10.1177/1367549409352277

Huron, D. (2006). *Sweet anticipation*. MIT Press.

Huron, D. (2012). *Aesthetics* (S. Hallam, I. Cross, & M. Thaut, Eds.; Vol. 1). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199298457.013.0014

Irsik, V. C., Johnsrude, I. S., & Herrmann, B. (2022). Neural Activity during Story Listening Is Synchronized across Individuals Despite Acoustic Masking. *Journal of Cognitive Neuroscience*, *34*(6), 933–950. https://doi.org/10.1162/jocn_a_01842

Isik, A. I., & Vessel, E. A. (2019). Continuous ratings of movie watching reveal idiosyncratic dynamics of aesthetic enjoyment. *PLOS ONE*, *14*(10), e0223896. https://doi.org/10.1371/journal.pone.0223896

Iverson, P., & Krumhansl, C. L. (1993). Isolating the dynamic attributes of musical timbre. *The Journal of the Acoustical Society of America*, *94*(5), 2595–2603. https://doi.org/10.1121/1.407371

Iwanaga, M., Kobayashi, A., & Kawasaki, C. (2005). Heart rate variability with repetitive exposure to music. *Biological Psychology*, *70*(1), 61–66. https://doi.org/10.1016/j.biopsycho.2004.11.015

Jacobsen, T. (2019). COMMENT Settings, Sensors, and Tasks: Comment on Scherer, Trznadel, Fantini, and Coutinho (2019). *Psychology of Aesthetics, Creativity, and the Arts*, *13*(3).

Jakubowski, K., Eerola, T., Alborno, P., Volpe, G., Camurri, A., & Clayton, M. (2017). Extracting Coarse Body Movements from Video in Music Performance: A Comparison of Automated Computer Vision Techniques with Motion Capture Data. *Frontiers in Digital Humanities*, *4*. https://doi.org/10.3389/fdigh.2017.00009

Janata, P., Tomic, S. T., & Haberman, J. M. (2012). Sensorimotor coupling in music and the psychology of the groove. *Journal of Experimental Psychology: General*, *141*(1), 54–75. https://doi.org/10.1037/a0024208

Jessen, S., & Kotz, S. A. (2011). The temporal dynamics of processing emotions from vocal, facial, and bodily expressions. *NeuroImage*, *58*(2), 665–674. https://doi.org/10.1016/j.neuroimage.2011.06.035

Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review*, *96*(3), 459–491. https://doi.org/10.1037/0033-295X.96.3.459

Juslin, P. N. (2013). From everyday emotions to aesthetic emotions: Towards a unified theory of musical emotions. *Physics of Life Reviews*, *10*(3), 235–266. https://doi.org/10.1016/j.plrev.2013.05.008

Juslin, P. N., & Laukka, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological Bulletin*, *129*(5), 770–814. https://doi.org/10.1037/0033-2909.129.5.770

Juslin, P. N., Liljeström, S., Västfjäll, D., Barradas, G., & Silva, A. (2008). An experience sampling study of emotional reactions to music: Listener, music, and situation. *Emotion*, *8*(5), 668–683. https://doi.org/10.1037/a0013505

Juslin, P. N., & Västfjäll, D. (2008). Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and Brain Sciences*, *31*(5), 559–575. https://doi.org/10.1017/S0140525X08005293

Kaneshiro, B., Nguyen, D. T., Norcia, A. M., Dmochowski, J. P., & Berger, J. (2020). Natural music evokes correlated EEG responses reflecting temporal structure and beat. *NeuroImage*, *214*, 116559. https://doi.org/10.1016/j.neuroimage.2020.116559

Kaneshiro, B., Nguyen, D. T., Norcia, A. M., Dmochowski, J. P., & Berger, J. (2021). *Inter-Subject EEG Correlation Reflects Time-Varying Engagement with Natural Music* [Preprint]. Neuroscience. https://doi.org/10.1101/2021.04.14.439913

Kang, O., & Wheatley, T. (2017). Pupil dilation patterns spontaneously synchronize across individuals during shared attention. *Journal of Experimental Psychology: General*, *146*(4), 569–576. https://doi.org/10.1037/xge0000271

Kappenman, E. S., & Luck, S. J. (2011). *ERP Components: The Ups and Downs of Brainwave Recordings*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195374148.013.0014

Karno, M., & Konečni, V. J. (1992). The Effects of Structural Interventions in the First Movement of Mozart's Symphony in G Minor K. 550 on Aesthetic Preference. *Music Perception*, *10*(1), 63–72. https://doi.org/10.2307/40285538

Kasdan, A. V., Burgess, A. N., Pizzagalli, F., Scartozzi, A., Chern, A., Kotz, S. A., Wilson, S. M., & Gordon, R. L. (2022). Identifying a brain network for musical rhythm: A functional neuroimaging meta-analysis and systematic review. *Neuroscience & Biobehavioral Reviews*, *136*, 104588. https://doi.org/10.1016/j.neubiorev.2022.104588

Kauppi. (2010). Inter-subject correlation of brain hemodynamic responses during watching a movie: Localization in space and frequency. *Frontiers in Neuroinformatics*. https://doi.org/10.3389/fninf.2010.00005

Kelly, J. (2023). BBC Singers: Music and the BBC – and beyond. *The Musicians' Company*. https://wcom.org.uk/news/bbc-singers-music-and-the-bbc-and-beyond/

Kern, P., Heilbron, M., de Lange, F. P., & Spaak, E. (2022). Cortical activity during naturalistic music listening reflects short-range predictions based on long-term experience. *ELife*, *11*, e80935. https://doi.org/10.7554/eLife.80935

Khalfa, S., Isabelle, P., Jean-Pierre, B., & Manon, R. (2002). Event-related skin conductance responses to musical emotions in humans. *Neuroscience Letters*, *328*(2), 145–149. https://doi.org/10.1016/S0304-3940(02)00462-7

Khalfa, S., Roy, M., Rainville, P., Dalla Bella, S., & Peretz, I. (2008). Role of tempo entrainment in psychophysiological differentiation of happy and sad music? *International Journal of Psychophysiology*, *68*(1), 17–26. https://doi.org/10.1016/j.ijpsycho.2007.12.001

Ki, J. J., Kelly, S. P., & Parra, L. C. (2016). Attention Strongly Modulates Reliability of Neural Responses to Naturalistic Narrative Stimuli. *Journal of Neuroscience*, *36*(10), 3092–3101. https://doi.org/10.1523/JNEUROSCI.2942-15.2016

Kirsch, L. P., Dawson, K., & Cross, E. S. (2015). Dance experience sculpts aesthetic perception and related brain circuits. *Annals of the New York Academy of Sciences*, *1337*(1), 130–139. https://doi.org/10.1111/nyas.12634

Kirsch, L. P., Drommelschmidt, K. A., & Cross, E. S. (2013). The impact of sensorimotor experience on affective evaluation of dance. *Frontiers in Human Neuroscience*, *7*. https://doi.org/10.3389/fnhum.2013.00521

Kirsch, L. P., Snagg, A., Heerey, E., & Cross, E. S. (2016). The Impact of Experience on Affective Responses during Action Observation. *PLOS ONE*, *11*(5), e0154681. https://doi.org/10.1371/journal.pone.0154681

Klimesch, W. (2018). The frequency architecture of brain and brain body oscillations: An analysis. *European Journal of Neuroscience*, *48*(7), 2431–2453. https://doi.org/10.1111/ejn.14192

Klimesch, W., Sauseng, P., & Hanslmayr, S. (2007). EEG alpha oscillations: The inhibition–timing hypothesis. *Brain Research Reviews*, *53*(1), 63–88. https://doi.org/10.1016/j.brainresrev.2006.06.003

Klucharev, V., Möttönen, R., & Sams, M. (2003). Electrophysiological indicators of phonetic and non-phonetic multisensory interactions during audiovisual speech perception. *Cognitive Brain Research*, *18*(1), 65–75. https://doi.org/10.1016/j.cogbrainres.2003.09.004

Koelsch, S. (2009). Music-syntactic processing and auditory memory: Similarities and differences between ERAN and MMN. *Psychophysiology*, *46*(1), 179–190. https://doi.org/10.1111/j.1469-8986.2008.00752.x

Koelsch, S. (2010). Towards a neural basis of music-evoked emotions. *Trends in Cognitive Sciences*, *14*(3), 131–137. https://doi.org/10.1016/j.tics.2010.01.002

Koelsch, S., Gunter, T., Friederici, A. D., & Schröger, E. (2000). Brain Indices of Music Processing: "Nonmusicians" are Musical. *Journal of Cognitive Neuroscience*, *12*(3), 520–541. https://doi.org/10.1162/089892900562183

Koelsch, S., & Jäncke, L. (2015). Music and the heart. *European Heart Journal*, *36*(44), 3043–3049. https://doi.org/10.1093/eurheartj/ehv430

Koelsch, S., Kilches, S., Steinbeis, N., & Schelinski, S. (2008). Effects of Unexpected Chords and of Performer's Expression on Brain Responses and Electrodermal Activity. *PLOS ONE*, *3*(7), e2631. https://doi.org/10.1371/journal.pone.0002631

Koelsch, S., Vuust, P., & Friston, K. (2019). Predictive Processes and the Peculiar Case of Music. *Trends in Cognitive Sciences*, *23*(1), 63–77. https://doi.org/10.1016/j.tics.2018.10.006

Kolb, B. (2000). You Call This Fun? Reactions of Young First-time Attendees to a Classical Concert. In D. Weidmann & S. Frederickson (Eds.), *Music Industry Issues and Studies* (Vol. 1, pp. 12–28). Music & Entertainment Industry Educators Association.

Kolb, B. M. (2001). The Decline of the Subscriber Base: A Study of the Philharmonia Orchestra Audience. *International Journal of Arts Management*, *3*(2), 51–59. https://doi.org/10.2307/41064723

Konečni, V. J. (2005). The aesthetic trinity: Awe, being moved, thrills. *Bulletin of Psychology and the Arts*, *5*(2), 27–44. https://doi.org/DOI: 10.1037/e674862010-005

Konvalinka, I., Xygalatas, D., Bulbulia, J., Schjødt, U., Jegindø, E.-M., Wallot, S., Van Orden, G., & Roepstorff, A. (2011). Synchronized arousal between performers and related spectators in a fire-walking ritual. *Proceedings of the National Academy of Sciences*, *108*(20), 8514–8519. https://doi.org/10.1073/pnas.1016955108

Kostoulas, T., Chanel, G., Muszynski, M., Lombardo, P., & Pun, T. (2015). Identifying aesthetic highlights in movies from clustering of physiological and behavioral signals. *2015 Seventh International Workshop on Quality of Multimedia Experience (QoMEX)*, 1–6. https://doi.org/10.1109/QoMEX.2015.7148098

Kostoulas, T., Chanel, G., Muszynski, M., Lombardo, P., & Pun, T. (2017). Films, Affective Computing and Aesthetic Experience: Identifying Emotional and Aesthetic Highlights from Multimodal Signals in a Social Setting. *Frontiers in ICT*, *4*, 11. https://doi.org/10.3389/fict.2017.00011

Kotz, S. A., Ravignani, A., & Fitch, W. T. (2018). The Evolution of Rhythm Processing. *Trends in Cognitive Sciences*, *22*(10), 896–910. https://doi.org/10.1016/j.tics.2018.08.002

Krause, A. E., North, A. C., & Hewitt, L. Y. (2015). Music-listening in everyday life: Devices and choice. *Psychology of Music*, *43*(2), 155–170. https://doi.org/10.1177/0305735613496860

Krumhansl, C. L. (1990). Tonal Hierarchies and Rare Intervals in Music Cognition. *Music Perception*, *7*(3), 309–324. https://doi.org/10.2307/40285467

Krumhansl, C. L. (1996). A Perceptual Analysis of Mozart's Piano Sonata K. 282: Segmentation, Tension, and Musical Ideas. *Music Perception*, *13*(3), 401–432. https://doi.org/10.2307/40286177

Krumhansl, C. L. (1997). An exploratory study of musical emotions and psychophysiology. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *51*(4), 336–353. https://doi.org/10.1037/1196-1961.51.4.336

Krumhansl, C. L. (2001). *Cognitive Foundations of Musical Pitch*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195148367.001.0001

Krumhansl, C. L., & Kessler, E. J. (1982). Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review*, *89*(4), 334–368. https://doi.org/10.1037/0033-295X.89.4.334

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). {lmerTest} Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Lahdelma, I., & Eerola, T. (2015). Theoretical Proposals on How Vertical Harmony May Convey Nostalgia and Longing in Music. *Empirical Musicology Review*, *10*(3), 245. https://doi.org/10.18061/emr.v10i3.4534

Lakatos, P., Shah, A. S., Knuth, K. H., Ulbert, I., Karmos, G., & Schroeder, C. E. (2005). An Oscillatory Hierarchy Controlling Neuronal Excitability and Stimulus Processing in the Auditory Cortex. *Journal of Neurophysiology*, *94*(3), 1904–1911. https://doi.org/10.1152/jn.00263.2005

Lamont, A. (2011). University students' strong experiences of music: Pleasure, engagement, and meaning. *Musicae Scientiae*, *15*(2), 229–249. https://doi.org/10.1177/1029864911403368

Lamont, A. (2012). Emotion, engagement and meaning in strong experiences of music performance. *Psychology of Music*, *40*(5), 574–594. https://doi.org/10.1177/0305735612448510

Lang, P. J., Greenwald, M. K., Bradley, M. M., & Hamm, A. O. (1993). Looking at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiology*, *30*(3), 261–273. https://doi.org/10.1111/j.1469-8986.1993.tb03352.x

Lange, E. B., & Frieler, K. (2018). Challenges and Opportunities of Predicting Musical Emotions with Perceptual and Automatized Features. *Music Perception*, *36*(2), 217–242. https://doi.org/10.1525/mp.2018.36.2.217

Lange, E. B., Fünderich, J., & Grimm, H. (2022). Multisensory integration of musical emotion perception in singing. *Psychological Research*. https://doi.org/10.1007/s00426-021-01637-9

Lange, E. B., Zweck, F., & Sinn, P. (2017). Microsaccade-rate indicates absorption by music listening. *Consciousness and Cognition*, *55*, 59–78. https://doi.org/10.1016/j.concog.2017.07.009

Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review*, *106*(1), 119–159. https://doi.org/10.1037/0033-295X.106.1.119

Large, E. W., & Snyder, J. S. (2009). Pulse and Meter as Neural Resonance. *Annals of the New York Academy of Sciences*, *1169*(1), 46–57. https://doi.org/10.1111/j.1749-6632.2009.04550.x

Larsen, J. T., Norris, C. J., & Cacioppo, J. T. (2003). Effects of positive and negative affect on electromyographic activity over zygomaticus major and corrugator supercilii. *Psychophysiology*, *40*(5), 776–785. https://doi.org/10.1111/1469-8986.00078

Lartillot, O., & Toiviainen, P. (2007). A Matlab toolbox for musical feature extraction from audio. *International Conference on Digital Audio Effects*, *237*.

Larue, J., Wolf, E. K., Bonds, M. E., Walsh, S., & Wilson, C. (2001). Symphony. In J. Larue, E. K. Wolf, M. E. Bonds, S. Walsh, & C. Wilson, *Oxford Music Online*. Oxford University Press. https://doi.org/10.1093/gmo/9781561592630.article.27254

Laurier, C., Lartillot, O., Eerola, T., & Toiviainen, P. (2009). Exploring relationships between audio features and emotion in music. *Frontiers in Human Neuroscience*, *3*. https://doi.org/10.3389/conf.neuro.09.2009.02.033

Leder, H., Belke, B., Oeberst, A., & Augustin, D. (2004). A model of aesthetic appreciation and aesthetic judgments. *British Journal of Psychology*, *95*(4), 489–508. https://doi.org/10.1348/0007126042369811

Leder, H., & Nadal, M. (2014). Ten years of a model of aesthetic appreciation and aesthetic judgments: The aesthetic episode - Developments and challenges in empirical aesthetics. *British Journal of Psychology*, *105*(4), 443–464. https://doi.org/10.1111/bjop.12084

Lenth, R. V. (2021). *emmeans: Estimated Marginal Means, aka Least-Squares Means*. https://CRAN.R-project.org/package=emmeans

Lerdahl, F., & Jackendoff, R. (1977). Toward a Formal Theory of Tonal Music. *Journal of Music Theory*, *21*(1), 111. https://doi.org/10.2307/843480

Lerdahl, F., & Jackendoff, R. (1983). An Overview of Hierarchical Structure in Music. *Music Perception: An Interdisciplinary Journal*, *1*(2), 229–252. https://doi.org/10.2307/40285257

Levinson, J. (2009). *Philosophical Aesthetics: An Overview* (J. Levinson, Ed.; Vol. 1). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199279456.003.0001

Lewis, M. (2011). Inside and Outside: The Relation between Emotional States and Expressions. *Emotion Review*, *3*(2), 189–196. https://doi.org/10.1177/1754073910387947

Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology*, *49*(4), 764–766. https://doi.org/10.1016/j.jesp.2013.03.013

Liaw, A., & Wiener, M. (2002). Classification and regression by random-Forest. *R News*, *2*(3), 18–22.

Lillywhite, A., Nijhof, D., Glowinski, D., Giordano, B. L., Camurri, A., Cross, I., & Pollick, F. E. (2022). A functional magnetic resonance imaging examination of audiovisual observation of a point-light string quartet using intersubject correlation and physical feature analysis. *Frontiers in Neuroscience*, *16*, 921489. https://doi.org/10.3389/fnins.2022.921489

Lindenberger, U., Li, S.-C., Gruber, W., & Müller, V. (2009). Brains swinging in concert: Cortical phase synchronization while playing guitar. *BMC Neuroscience*, *10*(1), 22. https://doi.org/10.1186/1471-2202-10-22

Locher, P., Overbeeke, K., & Wensveen, S. (2010). Aesthetic Interaction: A Framework. *Design Issues*, *26*(2). https://www.jstor.org/stable/20749943

Lokki, T., Pätynen, J., Kuusinen, A., & Tervo, S. (2012). Disentangling preference ratings of concert hall acoustics using subjective sensory profiles. *The Journal of the Acoustical Society of America*, *132*(5), 3148–3161. https://doi.org/10.1121/1.4756826

Lokki, T., Pätynen, J., Kuusinen, A., Vertanen, H., & Tervo, S. (2011). Concert hall acoustics assessment with individually elicited attributes. *The Journal of the Acoustical Society of America*, *130*(2), 835–849. https://doi.org/10.1121/1.3607422

Long, J. A. (2020). *jtools: Analysis and Presentation of Social Scientific Data* (R package version 2.1.0). https://cran.r-project.org/package=jtools

Lorig, T. S. (2016). The Respiratory System. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of Psychophysiology* (4th ed., pp. 244–257). Cambridge University Press. https://doi.org/10.1017/9781107415782.011

Loui, P., & Wessel, D. (2007). Harmonic expectation and affect in Western music: Effects of attention and training. *Perception & Psychophysics*, *69*(7), 1084–1092. https://doi.org/10.3758/BF03193946

Luck, G., Toiviainen, P., & Thompson, M. R. (2010). Perception of Expression in Conductors' Gestures: A Continuous Response Study. *Music Perception*, *28*(1), 47–57. https://doi.org/10.1525/mp.2010.28.1.47

Lundqvist, L.-O., Carlsson, F., Hilmersson, P., & Juslin, P. N. (2008). Emotional responses to music: Experience, expression, and physiology: *Psychology of Music*. https://doi.org/10.1177/0305735607086048

Luo, H., & Poeppel, D. (2007). Phase Patterns of Neuronal Responses Reliably Discriminate Speech in Human Auditory Cortex. *Neuron*, *54*(6), 1001–1010. https://doi.org/10.1016/j.neuron.2007.06.004

Lyytinen, H., Blomberg, A.-P., & Näätänen, R. (1992). Event-Related Potentials and Autonomic Responses to a Change in Unattended Auditory Stimuli. *Psychophysiology*, *29*(5), 523–534. https://doi.org/10.1111/j.1469-8986.1992.tb02025.x

Madison, G., & Schiölde, G. (2017). Repeated Listening Increases the Liking for Music Regardless of Its Complexity: Implications for the Appreciation and Aesthetics of Music. *Frontiers in Neuroscience*, *11*. https://doi.org/10.3389/fnins.2017.00147

Madsen, J., Margulis, E. H., Simchy-Gross, R., & Parra, L. C. (2019). Music synchronizes brainwaves across listeners with strong effects of repetition, familiarity and training. *Scientific Reports*, *9*(1), 3576. https://doi.org/10.1038/s41598-019-40254-w

Madsen, J., & Parra, L. C. (2022). Cognitive processing of a common stimulus synchronizes brains, hearts, and eyes. *PNAS Nexus*, *1*(1), pgac020. https://doi.org/10.1093/pnasnexus/pgac020

Magnee, M. J. C. M., Stekelenburg, J. J., Kemner, C., & de Gelder, B. (2007). Similar facial electromyographic responses to faces, voices, and body expressions: *NeuroReport*, *18*(4), 369–372. https://doi.org/10.1097/WNR.0b013e32801776e6

Margulis, E. H. (2014). *On repeat: How music plays the mind*. Oxford University Press.

Maslow, A. H. (1961). Peak experiences as acute identity experiences. *The American Journal of Psychoanalysis*, *21*(2), 254–262. https://doi.org/10.1007/BF01873126

Matsumoto, D. (1987). The role of facial response in the experience of emotion: More methodological problems and a meta-analysis. *Journal of Personality and Social Psychology*, *52*(4), 769–774. https://doi.org/10.1037/0022-3514.52.4.769

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language*, *94*, 305–315. https://doi.org/10.1016/j.jml.2017.01.001

McAdams, S., Vines, B. W., Vieillard, S., Smith, B. K., & Reynolds, R. (2004). Influences of large-scale form on continuous ratings in response to a contemporary piece in a live concert setting. *Music Perception*, *22*(2), 297–350. https://doi.org/10.1525/mp.2004.22.2.297

McAdams, S., Winsberg, S., Donnadieu, S., De Soete, G., & Krimphoff, J. (1995). Perceptual scaling of synthesized musical timbres: Common dimensions, specificities, and latent subject classes. *Psychological Research*, *58*(3), 177–192. https://doi.org/10.1007/BF00419633

Mencke, I., Omigie, D., Wald-Fuhrmann, M., & Brattico, E. (2019). Atonal Music: Can Uncertainty Lead to Pleasure? *Frontiers in Neuroscience*, *12*. https://www.frontiersin.org/articles/10.3389/fnins.2018.00979

Mencke, I., Quiroga-Martinez, D. R., Omigie, D., Michalareas, G., Schwarzacher, F., Haumann, N. T., Vuust, P., & Brattico, E. (2021). Prediction under uncertainty: Dissociating sensory from cognitive expectations in highly uncertain musical contexts. *Brain Research*, *1773*, 147664. https://doi.org/10.1016/j.brainres.2021.147664

Menninghaus, W., Schindler, I., Wagner, V., Wassiliwizky, E., Hanich, J., Jacobsen, T., & Koelsch, S. (2020). Aesthetic emotions are a key factor in aesthetic evaluation: Reply to Skov and Nadal (2020). *Psychological Review*, *127*(4), 650–654. https://doi.org/10.1037/rev0000213

Menninghaus, W., Wagner, V., Wassiliwizky, E., Schindler, I., Hanich, J., Jacobsen, T., & Koelsch, S. (2019). What are aesthetic emotions? *Psychological Review*, *126*(2), 171–195. https://doi.org/10.1037/rev0000135

Menon, V., & Levitin, D. J. (2005). The rewards of music listening: Response and physiological connectivity of the mesolimbic system. *NeuroImage*, *28*(1), 175–184. https://doi.org/10.1016/j.neuroimage.2005.05.053

Merrill, J., Ackermann, T.-I., & Czepiel, A. (2023). *The negative power of music: Effects of disliked music on psychophysiology* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/6escn

Merrill, J., Czepiel, A., Fink, L. T., Toelle, J., & Wald-Fuhrmann, M. (2021). The aesthetic experience of live concerts: Self-reports and psychophysiology. *Psychology of Aesthetics, Creativity, and the Arts*, *17*(2), 134–151. https://doi.org/10.1037/aca0000390

Minor, M. S., Wagner, T., Brewerton, F. J., & Hausman, A. (2004). Rock on! An elementary model of customer satisfaction with musical performances. *Journal of Services Marketing*, *18*(1), 7–18. https://doi.org/10.1108/08876040410520672

Moffat, D., Ronan, D., & Reiss, J. (2015). *An Evaluation of Audio Feature Extraction Toolboxes*. https://doi.org/10.13140/RG.2.1.1471.4640

Molholm, S., Ritter, W., Murray, M. M., Javitt, D. C., Schroeder, C. E., & Foxe, J. J. (2002). Multisensory auditory–visual interactions during early sensory processing in humans: A high-density electrical mapping study. *Cognitive Brain Research*, *14*(1), 115–128. https://doi.org/10.1016/S0926-6410(02)00066-6

Morrison & Selvey. (2014). The effect of conductor expressivity on choral ensemble evaluation. *Bulletin of the Council for Research in Music Education*, *199*, 7. https://doi.org/10.5406/bulcouresmusedu.199.0007

Mourot, L., Bouhaddi, M., Perrey, S., Rouillon, J.-D., & Regnard, J. (2004). Quantitative Poincar□ plot analysis of heart rate variability: Effect of endurance training. *European Journal of Applied Physiology*, *91*(1), 79–87. https://doi.org/10.1007/s00421-003-0917-0

Mulder, T., de Vries, S., & Zijlstra, S. (2005). Observation, imagination and execution of an effortful movement: More evidence for a central explanation of motor imagery. *Experimental Brain Research*, *163*(3), 344–351. https://doi.org/10.1007/s00221-004-2179-4

Müllensiefen, D., Gingras, B., Musil, J., & Stewart, L. (2014). Measuring the facets of musicality: The Goldsmiths Musical Sophistication Index (Gold-MSI). *Personality and Individual Differences*, *60*, S35. https://doi.org/10.1016/j.paid.2013.07.081

Müller, M. (2015). *Fundamentals of Music Processing: Audio, Analysis, Algorithms, Applications*. Springer International Publishing. https://doi.org/10.1007/978-3-319-21945-5

Müller, M., Höfel, L., Brattico, E., & Jacobsen, T. (2009). Electrophysiological Correlates of Aesthetic Music Processing: Comparing Experts with Laypersons. *Annals of the New York Academy of Sciences*, *1169*(1), 355–358. https://doi.org/10.1111/j.1749-6632.2009.04846.x

Muszynski, M., Kostoulas, T., Lombardo, P., Pun, T., & Chanel, G. (2018). Aesthetic Highlight Detection in Movies Based on Synchronization of Spectators' Reactions. *ACM Transactions on Multimedia Computing, Communications, and Applications*, *14*(3), 1–23. https://doi.org/10.1145/3175497

Mütze, H., Kopiez, R., & Wolf, A. (2020). The effect of a rhythmic pulse on the heart rate: Little evidence for rhythmical 'entrainment' and 'synchronization'. *Musicae Scientiae*, *24*(3), 377–400. https://doi.org/10.1177/1029864918817805

Näätänen, R., & Kreegipuu, K. (2011). *The Mismatch Negativity (MMN)*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195374148.013.0081

Nadal, M., & Vartanian, O. (2022). Empirical Aesthetics: An Overview. In M. Nadal & O. Vartanian (Eds.), *The Oxford Handbook of Empirical Aesthetics* (1st ed., pp. 3–38). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198824350.013.1

Nastase, S. A., Gazzola, V., Hasson, U., & Keysers, C. (2019). Measuring shared responses across subjects using intersubject correlation. *Social Cognitive and Affective Neuroscience*, *14*(6), 667–685. https://doi.org/10.1093/scan/nsz037

Nastase, S. A., Goldstein, A., & Hasson, U. (2020). Keep it real: Rethinking the primacy of experimental control in cognitive neuroscience. *NeuroImage*, *222*, 117254. https://doi.org/10.1016/j.neuroimage.2020.117254

Nelson, M. J., El Karoui, I., Giber, K., Yang, X., Cohen, L., Koopman, H., Cash, S. S., Naccache, L., Hale, J. T., Pallier, C., & Dehaene, S. (2017). Neurophysiological dynamics of phrase-structure building during sentence processing. *Proceedings of the National Academy of Sciences*, *114*(18). https://doi.org/10.1073/pnas.1701590114

Nemati, S., Akrami, H., Salehi, S., Esteky, H., & Moghimi, S. (2019). Lost in music: Neural signature of pleasure and its role in modulating attentional resources. *Brain Research*, *1711*, 7–15. https://doi.org/10.1016/j.brainres.2019.01.011

Nieminen, S., Istók, E., Brattico, E., Tervaniemi, M., & Huotilainen, M. (2011). The development of aesthetic responses to music and their underlying neural and psychological mechanisms. *Cortex*, *47*(9), 1138–1146. https://doi.org/10.1016/j.cortex.2011.05.008

North, A. C., & Hargreaves, D. J. (1995). Subjective complexity, familiarity, and liking for popular music. *Psychomusicology: A Journal of Research in Music Cognition*, *14*(1–2), 77–93. https://doi.org/10.1037/h0094090

North, A. C., Hargreaves, D. J., & Hargreaves, J. J. (2004). Uses of music in everyday life. *Music Perception*, *22*(1), 41–77. https://doi.org/10.1525/mp.2004.22.1.41

Obleser, J., Eisner, F., & Kotz, S. A. (2008). Bilateral Speech Comprehension Reflects Differential Sensitivity to Spectral and Temporal Features. *The Journal of Neuroscience*, *28*(32), 8116–8123. https://doi.org/10.1523/JNEUROSCI.1290-08.2008

Obleser, J., Herrmann, B., & Henry, M. J. (2012). Neural Oscillations in Speech: Don't be Enslaved by the Envelope. *Frontiers in Human Neuroscience*, *6*. https://doi.org/10.3389/fnhum.2012.00250

Obleser, J., & Weisz, N. (2012). Suppressed Alpha Oscillations Predict Intelligibility of Speech and its Acoustic Details. *Cerebral Cortex*, *22*(11), 2466–2477. https://doi.org/10.1093/cercor/bhr325

Olsen, K. N., Dean, R. T., & Stevens, C. J. (2014). A continuous measure of musical engagement contributes to prediction of perceived arousal and valence. *Psychomusicology: Music, Mind, and Brain*, *24*(2), 147–156. https://doi.org/10.1037/pmu0000044

Olsen, K. N., & Stevens, C. J. (2013). Psychophysiological Response to Acoustic Intensity Change in a Musical Chord. *Journal of Psychophysiology*, *27*(1), 16–26. https://doi.org/10.1027/0269-8803/a000082

Omigie, D., Frieler, K., Bär, C., Muralikrishnan, R., Wald-Fuhrmann, M., & Fischinger, T. (2021). Experiencing musical beauty: Emotional subtypes and their physiological and musico-acoustic correlates. *Psychology of Aesthetics, Creativity, and the Arts*, *15*(2), 197–215. https://doi.org/10.1037/aca0000271

O'Neill, K., & Egermann, H. (2022). Development of the Social Experience of a Concert Scales (SECS): The Social Experience of a Live Western Art Music Concert Influences People's Overall Enjoyment of an Event but not Their Emotional Response to the Music. *Music & Science*, *5*, 205920432211064. https://doi.org/10.1177/20592043221106478

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2011). FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. *Computational Intelligence and Neuroscience*, *2011*, 1–9. https://doi.org/10.1155/2011/156869

Orlandi, A., Cross, E. S., & Orgs, G. (2020). Timing is everything: Dance aesthetics depend on the complexity of movement kinematics. *Cognition*, *205*, 104446. https://doi.org/10.1016/j.cognition.2020.104446

Paccalin, C., & Jeannerod, M. (2000). Changes in breathing during observation of effortful actions. *Brain Research*, *862*(1–2), 194–200. https://doi.org/10.1016/S0006-8993(00)02145-4

Packer, J., & Ballantyne, J. (2011). The impact of music festival attendance on young people's psychological and social well-being. *Psychology of Music*, *39*(2), 164–181. https://doi.org/10.1177/0305735610372611

Page-Gould, E. (2016). Multilevel Modeling. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of Psychophysiology* (4th ed., pp. 662–678). Cambridge University Press. https://doi.org/10.1017/9781107415782.030

Palmer, C. (1997). MUSIC PERFORMANCE. *Annual Review of Psychology*, *48*(1), 115–138. https://doi.org/10.1146/annurev.psych.48.1.115

Palumbo, R. V., Marraccini, M. E., Weyandt, L. L., Wilder-Smith, O., McGee, H. A., Liu, S., & Goodwin, M. S. (2017). Interpersonal Autonomic Physiology: A Systematic Review of the Literature. *Personality and Social Psychology Review*, *21*(2), 99–141. https://doi.org/10.1177/1088868316628405

Panzarella, R. (1980). The Phenomenology of Aesthetic Peak Experiences. *Journal of Humanistic Psychology*, *20*(1), 69–85. https://doi.org/10.1177/002216788002000105

Park, H.-D., Correia, S., Ducorps, A., & Tallon-Baudry, C. (2014). Spontaneous fluctuations in neural responses to heartbeats predict visual detection. *Nature Neuroscience*, *17*(4), 612–618. https://doi.org/10.1038/nn.3671

Parviainen, T., Lyyra, P., & Nokia, M. S. (2022). Cardiorespiratory rhythms, brain oscillatory activity and cognition: Review of evidence and proposal for significance. *Neuroscience & Biobehavioral Reviews*, *142*, 104908. https://doi.org/10.1016/j.neubiorev.2022.104908

Patel, A. D. (2003). Language, music, syntax and the brain. *Nature Neuroscience*, *6*(7), 674–681. https://doi.org/10.1038/nn1082

Pearce, M. T. (2005). *The construction and evaluation of statistical models of melodic structure in music perception and composition.* Department of Computing, City University.

Peelle, J. E., & Davis, M. H. (2012). Neural Oscillations Carry Speech Rhythm through to Comprehension. *Frontiers in Psychology*, *3*. https://doi.org/10.3389/fpsyg.2012.00320

Peelle, J. E., Gross, J., & Davis, M. H. (2013). Phase-Locked Responses to Speech in Human Auditory Cortex are Enhanced During Comprehension. *Cerebral Cortex*, *23*(6), 1378–1387. https://doi.org/10.1093/cercor/bhs118

Pelowski, M., Forster, M., Tinio, P. P. L., Scholl, M., & Leder, H. (2017). Beyond the lab: An examination of key factors influencing interaction with 'real' and museum-based art. *Psychology of Aesthetics, Creativity, and the Arts*, *11*(3), 245–264. https://doi.org/10.1037/aca0000141

Pelowski, M., Markey, P. S., Lauring, J. O., & Leder, H. (2016). Visualizing the Impact of Art: An Update and Comparison of Current Psychological Models of Art Experience. *Frontiers in Human Neuroscience*, *10*. https://www.frontiersin.org/article/10.3389/fnhum.2016.00160

Pérez, P., Madsen, J., Banellis, L., Türker, B., Raimondo, F., Perlbarg, V., Valente, M., Niérat, M.-C., Puybasset, L., Naccache, L., Similowski, T., Cruse, D., Parra, L. C., & Sitt, J. D. (2021). Conscious processing of narrative stimuli synchronizes heart rate between individuals. *Cell Reports*, *36*(11), 109692. https://doi.org/10.1016/j.celrep.2021.109692

Phillips, M., Stewart, A. J., Wilcoxson, J. M., Jones, L. A., Howard, E., Willcox, P., Du Sautoy, M., & De Roure, D. (2020). What Determines the Perception of Segmentation in Contemporary Music? *Frontiers in Psychology*, *11*, 1001. https://doi.org/10.3389/fpsyg.2020.01001

Pines, R. (2002). *Surtitles* (Vol. 1). Oxford University Press. https://doi.org/10.1093/gmo/9781561592630.article.O007523

Piskorski, J., & Guzik, P. (2005). Filtering Poincaré plots. *Computational Methods in Science and Technology*, *11*(1), 39–48. https://doi.org/10.12921/CMST.2005.11.01.39-48

Pitts, S. E. (2004). 'Everybody wants to be Pavarotti': The Experience of Music for Performers and Audience at a Gilbert and Sullivan Festival. *Journal of the Royal Musical Association*, *129*(1), 143–160. https://doi.org/10.1093/jrma/fkh005

Pitts, S. E., & Spencer, C. P. (2008). Loyalty and Longevity in Audience Listening: Investigating Experiences of Attendance at a Chamber Music Festival. *Music and Letters*, *89*(2), 227–238. https://doi.org/10.1093/ml/gcm084

Platz, F., & Kopiez, R. (2012). When the Eye Listens: A Meta-analysis of How Audio-visual Presentation Enhances the Appreciation of Music Performance. *Music Perception*, *30*(1), 71–83. https://doi.org/10.1525/mp.2012.30.1.71

Plomp, R., & Levelt, W. J. M. (1965). Tonal Consonance and Critical Bandwidth. *The Journal of the Acoustical Society of America*, *38*(4), 548–560. https://doi.org/10.1121/1.1909741

Popescu, T., Widdess, R., & Rohrmeier, M. (2021). Western listeners detect boundary hierarchy in Indian music: A segmentation study. *Scientific Reports*, *11*(1), 3112. https://doi.org/10.1038/s41598-021-82629-y

Presicce, G., & Bailes, F. (2019). Engagement and visual imagery in music listening: An exploratory study. *Psychomusicology: Music, Mind, and Brain*, *29*(2–3), 136–155. https://doi.org/10.1037/pmu0000243

Proverbio, A. M., Manfrin, L., Arcari, L. A., De Benedetto, F., Gazzola, M., Guardamagna, M., Lozano Nasi, V., & Zani, A. (2015). Non-expert listeners show decreased heart rate and increased blood pressure (fear bradycardia) in response to atonal music. *Frontiers in Psychology*, *6*. https://www.frontiersin.org/articles/10.3389/fpsyg.2015.01646

Purves, D., & Williams, S. M. (Eds.). (2001). *Neuroscience* (2nd ed). Sinauer Associates.

R Core Team. (2021). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org/

R Development Core Team. (2020). *R: A Language and Environment for Statistical Computing* (Rstudio version 1.4.1106).

Randall, W. M., & Rickard, N. S. (2017). Personal Music Listening. *Music Perception*, *34*(5), 501–514. https://doi.org/10.1525/mp.2017.34.5.501

Reber, R., Schwarz, N., & Winkielman, P. (2004). Processing Fluency and Aesthetic Pleasure: Is Beauty in the Perceiver's Processing Experience? *Personality and Social Psychology Review*, *8*(4), 364–382. https://doi.org/10.1207/s15327957pspr0804_3

Revelle, W. (2022). *Package 'psych'*. https://personality-project.org/r/psych-manual.pdf

Revelle, W. (2021). *Package 'psych'*. https://personality-project.org/r/psych/

Richardson, D. C., Griffin, N. K., Zaki, L., Stephenson, A., Yan, J., Curry, T., Noble, R., Hogan, J., Skipper, J. I., & Devlin, J. T. (2020). Engagement in video and audio narratives: Contrasting self-report and physiological measures. *Scientific Reports*, *10*(1), 11298. https://doi.org/10.1038/s41598-020-68253-2

Richardson, M. J., Garcia, R. L., Frank, T. D., Gergor, M., & Marsh, K. L. (2012). Measuring group synchrony: A cluster-phase method for analyzing multivariate movement time-series. *Frontiers in Physiology*, *3*. https://doi.org/10.3389/fphys.2012.00405

Rickard, N. S. (2004). Intense emotional responses to music: A test of the physiological arousal hypothesis. *Psychology of Music*, *32*(4), 371–388. https://doi.org/10.1177/0305735604046096

Roose, H. (2008). Many-voiced or unisono? An inquiry into motives for attendance and aesthetic dispositions of the audience attending classical concerts. *Acta Sociologica*, *51*(3), 237–253. https://doi.org/10.1177/0001699308094168

Roy, M., Mailhot, J.-P., Gosselin, N., Paquette, S., & Peretz, I. (2009). Modulation of the startle reflex by pleasant and unpleasant music. *International Journal of Psychophysiology*, *71*(1), 37–42. https://doi.org/10.1016/j.ijpsycho.2008.07.010

RStudio Team. (2021). *RStudio: Integrated Development Environment for R*. RStudio, PBC. http://www.rstudio.com/

Ruiz-Padial, E., Vila, J., & Thayer, J. F. (2011). The effect of conscious and non-conscious presentation of biologically relevant emotion pictures on emotion modulated startle and phasic heart rate. *International Journal of Psychophysiology*, *79*(3), 341–346. https://doi.org/10.1016/j.ijpsycho.2010.12.001

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*(6), 1161–1178. https://doi.org/10.1037/h0077714

Sachs, M. E., Damasio, A., & Habibi, A. (2015). The pleasures of sad music: A systematic review. *Frontiers in Human Neuroscience*, *9*. https://doi.org/10.3389/fnhum.2015.00404

Sachs, M. E., Habibi, A., Damasio, A., & Kaplan, J. T. (2020). Dynamic intersubject neural synchronization reflects affective responses to sad music. *NeuroImage*, *218*, 116512. https://doi.org/10.1016/j.neuroimage.2019.116512

Sacks, O. (2006). The power of music. *Brain*, *129*(10), 2528–2532. https://doi.org/10.1093/brain/awl234

Salimpoor, V. N., Benovoy, M., Larcher, K., Dagher, A., & Zatorre, R. J. (2011). Anatomically distinct dopamine release during anticipation and experience of peak emotion to music. *Nature Neuroscience*, *14*(2), 257–262. https://doi.org/10.1038/nn.2726

Salimpoor, V. N., Benovoy, M., Longo, G., Cooperstock, J. R., & Zatorre, R. J. (2009). The Rewarding Aspects of Music Listening Are Related to Degree of Emotional Arousal. *PLOS ONE*, *4*(10), e7487. https://doi.org/10.1371/journal.pone.0007487

Sammler, D., Grigutsch, M., Fritz, T., & Koelsch, S. (2007). Music and emotion: Electrophysiological correlates of the processing of pleasant and unpleasant music. *Psychophysiology*, *44*(2), 293–304. https://doi.org/10.1111/j.1469-8986.2007.00497.x

Sarasso, P., Ronga, I., Pistis, A., Forte, E., Garbarini, F., Ricci, R., & Neppi-Modona, M. (2019). Aesthetic appreciation of musical intervals enhances behavioural and neurophysiological indexes of attentional engagement and motor inhibition. *Scientific Reports*, *9*(1), 18550. https://doi.org/10.1038/s41598-019-55131-9

Savage, P. E., Loui, P., Tarr, B., Schachner, A., Glowacki, L., Mithen, S., & Fitch, W. T. (2021). Music as a coevolved system for social bonding. *Behavioral and Brain Sciences*, *44*, e59. https://doi.org/10.1017/S0140525X20000333

Schaal, N. K., Bauer, A.-K. R., & Müllensiefen, D. (2014). Der Gold-MSI: Replikation und Validierung eines Fragebogeninstrumentes zur Messung *Musikalischer Erfahrenheit* anhand einer deutschen Stichprobe. *Musicae Scientiae*, *18*(4), 423–447. https://doi.org/10.1177/1029864914541851

Schaefer, T., Zimmermann, D., & Sedlmeier, P. (2014). How we remember the emotional intensity of past musical experiences. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00911

Schäfer, T., & Sedlmeier, P. (2009). From the functions of music to music preference. *Psychology of Music*, *37*(3), 279–300. https://doi.org/10.1177/0305735608097247

Scherer, K. R., Trznadel, S., Fantini, B., & Coutinho, E. (2019a). Assessing emotional experiences of opera spectators in situ. *Psychology of Aesthetics, Creativity, and the Arts*, *13*(3), 244–258. https://doi.org/10.1037/aca0000163

Scherer, K. R., Trznadel, S., Fantini, B., & Coutinho, E. (2019b). Comments on comments by Cupchik (2019) and Jacobsen (2019). *Psychology of Aesthetics, Creativity, and the Arts*, *13*(3), 264–265. https://doi.org/10.1037/aca0000246

Scherer, K. R., & Zentner, M. (2001). Emotional effects of music: Production rules. In *Music and emotion: Theory and research* (pp. 361–392). Oxford University Press.

Schindler, I., Hosoya, G., Menninghaus, W., Beermann, U., Wagner, V., Eid, M., & Scherer, K. R. (2017). Measuring aesthetic emotions: A review of the literature and a new assessment tool. *PLOS ONE*, *12*(6), e0178899. https://doi.org/10.1371/journal.pone.0178899

Schmälzle, R., Häcker, F. E. K., Honey, C. J., & Hasson, U. (2015). Engaged listeners: Shared neural processing of powerful political speeches. *Social Cognitive and Affective Neuroscience*, *10*(8), 1137–1143. https://doi.org/10.1093/scan/nsu168

Schön, D., & Morillon, B. (2019). Music and Language. In M. H. Thaut & D. A. Hodges (Eds.), *The Oxford Handbook of Music and the Brain* (pp. 390–416). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198804123.013.16

Schönberg, A. (1999). *Fundamentals of musical composition* (G. Strang & L. Stein, Eds.). faber and faber.

Schroeder, C. E., & Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. *Trends in Neurosciences*, *32*(1), 9–18. https://doi.org/10.1016/j.tins.2008.09.012

Schubert, E. (2003). Update of the Hevner Adjective Checklist. *Perceptual and Motor Skills*, *96*(3_suppl), 1117–1122. https://doi.org/10.2466/pms.2003.96.3c.1117

Schubert, E., Vincs, K., & Stevens, C. J. (2013). Identifying Regions of Good Agreement among Responders in Engagement with a Piece of Live Dance. *Empirical Studies of the Arts*, *31*(1), 1–20. https://doi.org/10.2190/EM.31.1.a

Schultz, B. G., Brown, R. M., & Kotz, S. A. (2021). Dynamic acoustic salience evokes motor responses. *Cortex*, *134*, 320–332. https://doi.org/10.1016/j.cortex.2020.10.019

Schutz, M., & Kubovy, M. (2009). Causality and cross-modal integration. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(6), 1791–1810. https://doi.org/10.1037/a0016455

Schutz, M., & Lipscomb, S. (2007). Hearing Gestures, Seeing Music: Vision Influences Perceived Tone Duration. *Perception*, *36*(6), 888–897. https://doi.org/10.1068/p5635

Seibert, C., Toelle, J., & Wald-Fuhrmann, M. (2020). The classical concert as an object of empirical aesthetics. In *Classical concert studies: A companion to contemporary research and performance* (pp. 351–360). Routledge.

Sethares, W. A. (1998). Consonance-Based Spectral Mappings. *Computer Music Journal*, *22*(1), 56. https://doi.org/10.2307/3681045

Shaffer, F., & Ginsberg, J. P. (2017). An Overview of Heart Rate Variability Metrics and Norms. *Frontiers in Public Health*, *5*. https://www.frontiersin.org/article/10.3389/fpubh.2017.00258

Sherman, N. (2022, November 18). Taylor Swift public ticket sale cancelled amid high demand. *BBC News*. https://www.bbc.com/news/business-63652665

Shusterman, R. (1997). The End of Aesthetic Experience. *The Journal of Aesthetics and Art Criticism*, *55*(1), 29. https://doi.org/10.2307/431602

Sidle, D. A. T., & Heron, P. A. (1976). Effects of Length of Training and Amount of Tone Frequency Change on Amplitude of Autonomic Components of the Orienting Response. *Psychophysiology*, *13*(4), 281–287. https://doi.org/10.1111/j.1469-8986.1976.tb03076.x

Silbert, L. J., Honey, C. J., Simony, E., Poeppel, D., & Hasson, U. (2014). Coupled neural systems underlie the production and comprehension of naturalistic narrative speech. *Proceedings of the National Academy of Sciences*, *111*(43). https://doi.org/10.1073/pnas.1323812111

Simony, E., Honey, C. J., Chen, J., Lositsky, O., Yeshurun, Y., Wiesel, A., & Hasson, U. (2016). Dynamic reconfiguration of the default mode network during narrative comprehension. *Nature Communications*, *7*(1), 12141. https://doi.org/10.1038/ncomms12141

Skov, M., & Nadal, M. (2020). There are no aesthetic emotions: Comment on Menninghaus et al. (2019). *Psychological Review*, *127*(4), 640–649. https://doi.org/10.1037/rev0000187

Skov, M., & Nadal, M. (Eds.). (2023). *The Routledge international handbook of neuroaesthetics*. Routledge, Taylor & Francis Group. https://doi.org/10.4324/9781003008675

Sloboda, J. A. (1991). Music Structure and Emotional Response: Some Empirical Findings. *Psychology of Music*, *19*(2), 110–120. https://doi.org/10.1177/0305735691192002

Sloboda, J. A. (2010). Music in everyday life: The role of emotions. In *Handbook of Music and Emotion: Theory, Research, Applications Handbook of Music and Emotion: Theory, Research, Applications* (pp. 492–514). Oxford University Press, Oxford. https://academic.oup.com/book/38621/chapter/335188249

Sloboda, J. A., Lamont, A., & Greasley, A. (2012). *Choosing to hear music* (S. Hallam, I. Cross, & M. Thaut, Eds.; Vol. 1). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199298457.013.0040

Sloboda, J. A., & O'Neill, S. A. (2001). Emotions in everyday listening to music. In *Music and emotion: Theory and research* (pp. 415–429). Oxford University Press.

Small, C. (1998). *Musicking: The meanings of performing and listening.*

Sosnowski, M., Clark, E., Latif, S., Macfarlane, P. W., & Tendera, M. (2005). Heart Rate Variability Fraction-A New Reportable Measure of 24-Hour R-R Interval Variation. *Annals of Noninvasive Electrocardiology*, *10*(1), 7–15. https://doi.org/10.1111/j.1542-474X.2005.00579.x

Steinbeis, N., Koelsch, S., & Sloboda, J. A. (2006). The Role of Harmonic Expectancy Violations in Musical Emotions: Evidence from Subjective, Physiological, and Neural Responses. *Journal of Cognitive Neuroscience*, *18*(8), 1380–1393. https://doi.org/10.1162/jocn.2006.18.8.1380

Stekelenburg, J., & Boxtel, A. V. (2001). Inhibition of pericranial muscle activity, respiration, and heart rate enhances auditory sensitivity. *Psychophysiology*, *38*(4), 629–641. https://doi.org/10.1111/1469-8986.3840629

Stekelenburg, J., & van Boxtel, A. (2002). Pericranial muscular, respiratory, and heart rate components of the orienting response. *Psychophysiology*, *39*(6), 707–722.

Stekelenburg, J., & Vroomen, J. (2007). Neural Correlates of Multisensory Integration of Ecologically Valid Audiovisual Events. *Journal of Cognitive Neuroscience*, *19*(12), 1964–1973. https://doi.org/10.1162/jocn.2007.19.12.1964

Swarbrick, D., Bosnyak, D., Livingstone, S. R., Bansal, J., Marsh-Rollo, S., Woolhouse, M. H., & Trainor, L. J. (2019). How Live Music Moves Us: Head Movement Differences in Audiences to Live Versus Recorded Music. *Frontiers in Psychology*, *9*, 2682. https://doi.org/10.3389/fpsyg.2018.02682

Szpunar, K. K., Schellenberg, E. G., & Pliner, P. (2004). Liking and Memory for Musical Stimuli as a Function of Exposure. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(2), 370–381. https://doi.org/10.1037/0278-7393.30.2.370

Tallon-Baudry, C. (1999). Oscillatory gamma activity in humans and its role in object representation. *Trends in Cognitive Sciences*, *3*(4), 151–162. https://doi.org/10.1016/S1364-6613(99)01299-1

Taruffi, L., Pehrs, C., Skouras, S., & Koelsch, S. (2017). Effects of Sad and Happy Music on Mind-Wandering and the Default Mode Network. *Scientific Reports*, *7*(1), 14396. https://doi.org/10.1038/s41598-017-14849-0

Task Force of the European Society of Cardiology, the N. A. S. of P. E. (1996). Heart Rate Variability: Standards of Measurement, Physiological Interpretation, and Clinical Use. *Circulation*, *93*(5), 1043–1065. https://doi.org/10.1161/01.CIR.93.5.1043

Tassinary, L. G., Cacioppo, J. T., & Vanman, E. J. (2016). The Somatic System. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of Psychophysiology* (4th ed., pp. 151–182). Cambridge University Press. https://doi.org/10.1017/9781107415782.008

Tervaniemi, M. (2023). The neuroscience of music – towards ecological validity. *Trends in Neurosciences*, *46*(5), 355–364. https://doi.org/10.1016/j.tins.2023.03.001

Tervaniemi, M., Pousi, S., Seppälä, M., & Makkonen, T. (2021). Brain oscillation recordings of the audience in a live concert-like setting. *Cognitive Processing*. https://doi.org/10.1007/s10339-021-01072-z

Thompson. (2006). Audience responses to a live orchestral concert. *Musicae Scientiae*, *10*(2), 215–244. https://doi.org/10.1177/102986490601000204

Thompson, Bullot, N. J., & Margulis, E. H. (2023). The psychological basis of music appreciation: Structure, self, source. *Psychological Review*, *130*(1), 260–284. https://doi.org/10.1037/rev0000364

Thompson, & Luck, G. (2012). Exploring relationships between pianists' body movements, their expressive intentions, and structural elements of the music. *Musicae Scientiae*, *16*(1), 19–40. https://doi.org/10.1177/1029864911423457

Thorau, C., & Ziemer, H. (Eds.). (2019). *The Oxford handbook of music listening in the nineteenth and twentieth centuries.* Oxford University Press.

Ticini, L. F., Urgesi, C., & Calvo-Merino, B. (2015). Embodied Aesthetics: Insight from Cognitive Neuroscience of Performing Arts. In A. Scarinzi (Ed.), *Aesthetics and the Embodied Mind: Beyond Art Theory and the Cartesian Mind-Body Dichotomy* (Vol. 73, pp. 103–115). Springer Netherlands. https://doi.org/10.1007/978-94-017-9379-7_7

Tillmann, B., & Bigand, E. (2004). The Relative Importance of Local and Global Structures in Music Perception. *The Journal of Aesthetics and Art Criticism*, *62*(2), 211–222. JSTOR.

Tillmann, B., Bigand, E., & Madurell, F. (1998). Local versus global processing of harmonic cadences in the solution of musical puzzles. *Psychological Research*, *61*(3), 157–174. https://doi.org/10.1007/s004260050022

Timmers, R., Marolt, M., Camurri, A., & Volpe, G. (2006). Listeners' emotional engagement with performances of a Scriabin étude: An explorative case study. *Psychology of Music*, *34*(4), 481–510. https://doi.org/10.1177/0305735606067165

Toelle, J., & Sloboda, J. A. (2021). *The audience as artist?* The audience's experience of participatory music. *Musicae Scientiae*, *25*(1), 67–91. https://doi.org/10.1177/1029864919844804

Toiviainen, P., & Carlson, E. (2022). Embodied Meter Revisited. *Music Perception*, *39*(3), 249–267. https://doi.org/10.1525/mp.2022.39.3.249

Toiviainen, P., & Krumhansl, C. L. (2003). Measuring and Modeling Real-Time Responses to Music: The Dynamics of Tonality Induction. *Perception*, *32*(6), 741–766. https://doi.org/10.1068/p3312

Toiviainen, P., Luck, G., & Thompson, M. R. (2010). Embodied Meter: Hierarchical Eigenmodes in Music-Induced Movement. *Music Perception*, *28*(1), 59–70. https://doi.org/10.1525/mp.2010.28.1.59

Tramo, M. J., Cariani, P. A., Delgutte, B., & Braida, L. D. (2001). Neurobiological Foundations for the Theory of Harmony in Western Tonal Music. *Annals of the New York Academy of Sciences*, *930*(1), 92–116. https://doi.org/10.1111/j.1749-6632.2001.tb05727.x

Tröndle, M., Greenwood, S., Ramakrishnan, C., Uhde, F., Egermann, H., & Tschacher, W. (2022). Integrated Methods: A Call for Integrative and Interdisciplinary Aesthetics Research. In M. Nadal & O. Vartanian (Eds.), *The Oxford Handbook of Empirical Aesthetics* (1st ed., pp. 359–382). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198824350.013.18

Trost, W., Ethofer, T., Zentner, M., & Vuilleumier, P. (2012). Mapping Aesthetic Musical Emotions in the Brain. *Cerebral Cortex*, *22*(12), 2769–2783. https://doi.org/10.1093/cercor/bhr353

Tsay, C.-J. (2013). Sight over sound in the judgment of music performance. *Proceedings of the National Academy of Sciences*, *110*(36), 14580–14585. https://doi.org/10.1073/pnas.1221454110

Tschacher, W., Greenwood, S., Egermann, H., Wald-Fuhrmann, M., Czepiel, A., Tröndle, M., & Meier, D. (2021). Physiological synchrony in audiences of live concerts. *Psychology of Aesthetics, Creativity, and the Arts*, *17*(2), 152–162. https://doi.org/10.1037/aca0000431

Tschacher, W., Greenwood, S., Kirchberg, V., Wintzerith, S., van den Berg, K., & Tröndle, M. (2012). Physiological correlates of aesthetic perception of artworks in a museum. *Psychology of Aesthetics, Creativity, and the Arts*, *6*(1), 96–103. https://doi.org/10.1037/a0023845

Tzanetakis, G., & Cook, P. (2002). Musical genre classification of audio signals. *IEEE Transactions on Speech and Audio Processing*, *10*(5), 293–302. https://doi.org/10.1109/TSA.2002.800560

van den Bosch, I., Salimpoor, V. N., & Zatorre, R. J. (2013). Familiarity mediates the relationship between emotional arousal and pleasure during music listening. *Frontiers in Human Neuroscience*, *7*. https://doi.org/10.3389/fnhum.2013.00534

van der Zwaag, M. D., Westerink, J. H. D. M., & van den Broek, E. L. (2011). Emotional and psychophysiological responses to tempo, mode, and percussiveness. *Musicae Scientiae*, *15*(2), 250–269. https://doi.org/10.1177/1029864911403364

van Wassenhove, V., Grant, K. W., & Poeppel, D. (2005). Visual speech speeds up the neural processing of auditory speech. *Proceedings of the National Academy of Sciences*, *102*(4), 1181–1186. https://doi.org/10.1073/pnas.0408949102

Van Zijl, A. G. W., & Luck, G. (2013). Moved through music: The effect of experienced emotions on performers' movement characteristics. *Psychology of Music*, *41*(2), 175–197. https://doi.org/10.1177/0305735612458334

Vanden Bosch der Nederlanden, C. M., Joanisse, M. F., & Grahn, J. A. (2020). Music as a scaffold for listening to speech: Better neural phase-locking to song than speech. *NeuroImage*, *214*, 116767. https://doi.org/10.1016/j.neuroimage.2020.116767

Vessel, E. A., & Rubin, N. (2010). Beauty and the beholder: Highly individual taste for abstract, but not real-world images. *Journal of Vision*, *10*(2), 1–14. https://doi.org/10.1167/10.2.18

Vickhoff, B., Malmgren, H., Åström, R., Nyberg, G., Engvall, M., Snygg, J., Nilsson, M., & Jörnsten, R. (2013). Music structure determines heart rate variability of singers. *Frontiers in Psychology*, *4*. https://www.frontiersin.org/articles/10.3389/fpsyg.2013.00334

Vines, B. W., Krumhansl, C. L., Wanderley, M. M., Dalca, I. M., & Levitin, D. J. (2011). Music to my eyes: Cross-modal interactions in the perception of emotions in musical performance. *Cognition*, *118*(2), 157–170. https://doi.org/10.1016/j.cognition.2010.11.010

Vines, B. W., Krumhansl, C. L., Wanderley, M. M., & Levitin, D. J. (2006). Cross-modal interactions in the perception of musical performance. *Cognition*, *101*(1), 80–113. https://doi.org/10.1016/j.cognition.2005.09.003

Vines, B. W., Wanderley, M. M., Krumhansl, C. L., Nuzzo, R. L., & Levitin, D. J. (2004). Performance Gestures of Musicians: What Structural and Emotional Information Do They Convey? In A. Camurri & G. Volpe (Eds.), *Gesture-Based Communication in Human-Computer Interaction* (pp. 468–478). Springer. https://doi.org/10.1007/978-3-540-24598-8_43

Voloh, B., & Womelsdorf, T. (2016). A Role of Phase-Resetting in Coordinating Large Scale Neural Networks During Attention and Goal-Directed Behavior. *Frontiers in Systems Neuroscience*, *10*. https://doi.org/10.3389/fnsys.2016.00018

Vrana, S. R. (1993). The psychophysiology of disgust: Differentiating negative emotional contexts with facial EMG. *Psychophysiology*, *30*(3), 279–286. https://doi.org/10.1111/j.1469-8986.1993.tb03354.x

Vroegh, T. P. (2018). *The pleasures of getting into the music: Absorption, and its role in the aesthetic appreciation of music* [Doctoralthesis]. Universitätsbibliothek Johann Christian Senckenberg.

Vroegh, T. P. (2019). Zoning-in or tuning-in? Identifying distinct absorption states in response to music. *Psychomusicology: Music, Mind, and Brain*, *29*(2–3), 156–170. https://doi.org/10.1037/pmu0000241

Vuoskoski, J. K., & Eerola, T. (2011). Measuring music-induced emotion: A comparison of emotion models, personality biases, and intensity of experiences. *Musicae Scientiae*, *15*(2), 159–173. https://doi.org/10.1177/1029864911403367

Vuoskoski, J. K., Gatti, E., Spence, C., & Clarke, E. F. (2016). Do visual cues intensify the emotional responses evoked by musical performance? A psychophysiological investigation. *Psychomusicology: Music, Mind, and Brain*, *26*(2), 179–188. http://dx.doi.org/10.1037/pmu0000142

Vuoskoski, J. K., Thompson, M. R., Clarke, E. F., & Spence, C. (2014). Crossmodal interactions in the perception of expressivity in musical performance. *Attention, Perception, & Psychophysics*, *76*(2), 591–604. https://doi.org/10.3758/s13414-013-0582-2

Vuust, P., Heggli, O. A., Friston, K. J., & Kringelbach, M. L. (2022). Music in the brain. *Nature Reviews Neuroscience*, *23*(5), 287–305. https://doi.org/10.1038/s41583-022-00578-5

Waddell, G., & Williamon, A. (2017). Eye of the Beholder: Stage Entrance Behavior and Facial Expression Affect Continuous Quality Ratings in Music Performance. *Frontiers in Psychology*, *8*, 513. https://doi.org/10.3389/fpsyg.2017.00513

Wald-Fuhrmann, M., Egermann, H., Czepiel, A., O'Neill, K., Weining, C., Meier, D., Tschacher, W., Uhde, F., Toelle, J., & Tröndle, M. (2021). Music Listening in Classical Concerts: Theory, Literature Review, and Research Program. *Frontiers in Psychology*, *12*. https://www.frontiersin.org/article/10.3389/fpsyg.2021.638783

Wanderley, M. M., Vines, B. W., Middleton, N., McKay, C., & Hatch, W. (2005). The Musical Significance of Clarinetists' Ancillary Gestures: An Exploration of the Field. *Journal of New Music Research*, *34*(1), 97–113. https://doi.org/10.1080/09298210500124208

Wassiliwizky, E., Koelsch, S., Wagner, V., Jacobsen, T., & Menninghaus, W. (2017). The emotional power of poetry: Neural circuitry, psychophysiology and compositional principles. *Social Cognitive and Affective Neuroscience*, *12*(8), 1229–1240. https://doi.org/10.1093/scan/nsx069

Weber, W. (2001). *Concert (ii)* (Vol. 1). Oxford University Press. https://doi.org/10.1093/gmo/9781561592630.article.06240

Weber, W. (2003). Consequences of Canon: The Institutionalization of Enmity between Contemporary and Classical Music. *Common Knowledge*, *9*(1), 78–99.

Webster, J. (1991). *Haydn's 'Farewell' Symphony and the Idea of Classical Style: Through-Composition and Cyclic Integration in his Instrumental Music* (1st ed.). Cambridge University Press. https://doi.org/10.1017/CBO9780511552434

Weineck, K., Wen, O. X., & Henry, M. J. (2022). Neural synchronization is strongest to the spectral flux of slow music and depends on familiarity and beat salience. *ELife*, *11*, e75515. https://doi.org/10.7554/eLife.75515

Weingarten, G. (2007, April 8). Pearls Before Breakfast: Can one of the nation's great musicians cut through the fog of a D.C. rush hour? Let's find out. *Washington Post*. https://www.washingtonpost.com/lifestyle/magazine/pearls-before-breakfast-can-one-of-the-nations-great-musicians-cut-through-the-fog-of-a-dc-rush-hour-lets-find-out/2014/09/23/8a6d46da-4331-11e4-b47c-f5889e061e5f_story.html

Williams, J. A., Margulis, E. H., Nastase, S. A., Chen, J., Hasson, U., Norman, K. A., & Baldassano, C. (2022). High-Order Areas and Auditory Cortex Both Represent the High-Level Event Structure of Music. *Journal of Cognitive Neuroscience*, *34*(4), 699–714. https://doi.org/10.1162/jocn_a_01815

Wilson, C. V., & Aiken, L. S. (1977). The Effect of Intensity Levels Upon Physiological and Subjective Affective Response to Rock Music. *Journal of Music Therapy*, *14*(2), 60–76. https://doi.org/10.1093/jmt/14.2.60

Wilson, S. M., Molnar-Szakacs, I., & Iacoboni, M. (2008). Beyond Superior Temporal Cortex: Intersubject Correlations in Narrative Speech Comprehension. *Cerebral Cortex*, *18*(1), 230–242. https://doi.org/10.1093/cercor/bhm049

Wingenbach, T. S. H., Brosnan, M., Pfaltz, M. C., Peyk, P., & Ashwin, C. (2020). Perception of Discrete Emotions in Others: Evidence for Distinct Facial Mimicry Patterns. *Scientific Reports*, *10*(1), 4692. https://doi.org/10.1038/s41598-020-61563-5

Winter, B. (2013). *Linear models and linear mixed effects models in R with linguistic applications* (arXiv:1308.5499). arXiv. http://arxiv.org/abs/1308.5499

Witvliet, C. V. O., & Vrana, S. R. (2007). Play it again Sam: Repeated exposure to emotionally evocative music polarises liking and smiling responses, and influences other affective reports, facial EMG, and heart rate. *Cognition & Emotion*, *21*(1), 3–25. https://doi.org/10.1080/02699930601000672

Wu, B., Horner, A., & Lee, C. (2014). *Musical timbre and emotion: The identification of salient timbral features in sustained musical instrument tones equalized in attack time and spectral centroid*. 928–934.

Zatorre, R. J., Chen, J. L., & Penhune, V. B. (2007). When the brain plays music: Auditory–motor interactions in music perception and production. *Nature Reviews Neuroscience*, *8*(7), 547–558. https://doi.org/10.1038/nrn2152

Zbikowski, L. M. (2002). *Conceptualizing music: Cognitive structure, theory, and analysis*. Oxford university press.

Zeki, S. (1999). Art and the brain. *Journal of Consciousness Studies*, *6*(6–7), 76–96.

Zentner, M., Grandjean, D., & Scherer, K. R. (2008). Emotions evoked by the sound of music: Characterization, classification, and measurement. *Emotion*, *8*(4), 494–521. https://doi.org/10.1037/1528-3542.8.4.494

Appendix

Impact Paragraph

Live music listening is a socially and culturally important human activity. We choose to invest time and money into this activity, as it enriches our lives emotionally, intellectually, and socially. Prior to 2020, attendance of live music concerts was generally increasing in countries like the U.K. (Art Council UK, 2014/2015), Germany (Deutscher Bühnenverein, 2022), and the Netherlands (Grace, 2020). The importance of live music has been highlighted by the lockdowns that occurred during the COVID-19 pandemic demonstrated how powerful music is for promoting social connectedness during a seemingly dark time (Fink et al., 2021). Decisions to substantially cut funding to nationally important institutions such as the English National Opera (Connolly, 2022) and BBC (British Broadcasting Corporation) Singers (Kelly, 2023) have sparked intense criticism from musicians, scholars, and the general public. As this illustrates the social and cultural values of live music, the goal of the current research was to increase our understanding of live music listening experiences, specifically in Western classical music concerts. To assess the concert experience, the current research used self-report questionnaires and peripheral physiological measures. The latter consisted of fight-and-flight system responses (which typically reflect states of attention and emotional arousal) and facial muscle activation (frowning and smiling). How our findings relate to wellbeing and inform future research is discussed below.

We know from previous studies that attending – as well as performing in – live music events can positively impact wellbeing, by evoking positive aesthetic experiences (Gabrielsson & Wik, 2003; Lamont, 2011). A main finding of the current thesis was that concerts indeed evoke high liking and positive emotions, despite occasional negative emotions. Further, this positive experience was enhanced in live, audio-visual music performances, compared to audio-only recordings. This suggests that certain components of concert programs together with their visual aspects might be key elements evoking positive experiences. In turn, these positive experiences likely benefit overall wellbeing. By assessing (synchrony of) physiological responses, this thesis also highlighted potential mechanisms of how such positive feelings are evoked, that is, through engagement, emotional contagion (mirroring of emotion), and motor mimicry. Although the current research was conducted with generally healthy participants, such understanding of positive experiences could be extended in future to assess whether this positive impact on wellbeing could also be beneficial for general (physical and mental) health.

Concerts have important social benefits. Indeed, social bonding is one evolutionary explanation of music (Kotz et al., 2018; Savage et al., 2021). Live music performances create important social experiences

(Dearn & Price, 2016; Packer & Ballantyne, 2011). They also provide a sense of togetherness and community belonging (Pitts, 2004). The collective experience was explored in the current thesis by calculating time-domain and phase-domain physiological synchrony in audiences (Chapter 3 and 5). These results highlight that music synchronises physiological responses of audience members, though time-domain measures were more robust than phase synchrony measures. For example, music synchronised heart rate acceleration and deceleration across audience members, but not their actual heart beats. One reason for the lack of phase synchrony is explained by the customary limited physical movement and interaction of the audience in Western classical concerts. Other genres, such as pop and rock, greater interaction through dancing together and singing along to the music is not only allowed, but is a positive part of the experience. In these genres, perhaps phase synchrony measure is more relevant. Although this thesis did not examine the social aspect directly, future studies could use similar paradigms and analysis methods to further evaluate and understand mechanisms of social bonding during music listening. The synchrony analysis pipelines could be applied to other kinds of signals such as brain signals, eye tracking, and gastric rhythms, to further inform brain-body interactions. Such measures can be extended not only to music listening and performance, but to other situations, such as educational settings (e.g., synchrony between teacher and students) as well as social activities like conversations, dance, and team sport.

The results of this current research could inform those who plan and curate concerts. Western classical concerts are generally perceived as being very formal events, which do not always encourage attendance. In my own concert management work (see Appendix, C.V.), I was able to experience the great effort concert organisers put into diversifying their audiences by careful concert program planning, by incorporating educational aspects, and by increasing audience interaction. This thesis extends our empirical understanding of aspects that might make concerts more appealing, for example, a balanced program of familiar music, mixed with new, potentially slightly challenging, music. This and future investigations could shed light on potential areas of focus for marketing concerts to wider audiences. Such investments are vital, as concerts benefit society not only culturally, but also economically. Successful concerts and music festivals have increased visibility and tourism in smaller cities and places, for example Darmstadt (known for the New Music), Bayreuth (known for the world-famous Wagner festival), and Aldeburgh (known for its innovative music festival). More concert goers would increase the chance of investments being made to promote music more widely. This might subsequently sustain funding necessary to make music more affordable, which in turn benefits society more generally.

In summary, concert research and attendance to concerts therefore are not only valuable for positive wellbeing and social benefit of individuals, but are also of substantial societal and cultural value.

In support of accessibility of scientific knowledge, the studies in Chapter 2-5 have been or are planned to be made publicly available in preprints (prior to paper acceptance) and open access journals. This thesis also contributes to Open Science, i.e., practices that make science available, accessible, reusable, and transparent. The piano concert data as well as the analysis code has been uploaded to the Open Science Framework. We note the repository is currently private while the paper is in review, but will be made public once the paper has been accepted. The research in this thesis has been presented as a keynote presentation as well as being presented in both posters and talks at other international scientific conferences. These presentations contribute to the growing scientific field of body-brain understanding (Criscuolo et al., 2022) as well as the advancement of neuroscience in real-world settings (Tervaniemi, 2023).

Appendix

Curriculum Vitae

Appendix: C.V.

Anna Czepiel was born on June 26, 1995 in Oxford, England. She completed her secondary schooling at The Cherwell School, in Oxford. She gained a first class (honours) bachelor's degree in music at the University of York, England, in 2016. Her Bachelor thesis explored (sub)conscious gestures in piano players. During this time, Anna accumulated experience in concert management: working with the Oxford Philharmonic (2012-2018), being a concert manager for the University of York Music Society and The Chimera Ensemble (2014-2015). She was also on the planning committee for the York Spring Festival of New Music (2015-2016), and was part of the Hesse Student Scheme at the Aldeburgh Festival (2015). Anna continued her education to a Master's degree in Music, Mind, and Technology with a minor in Cognitive Neuroscience at the University of Jyväskylä, Finland. She conducted her Master thesis on the motion features of piano performances. Inspired by the exciting research on music and language neuroscience encountered at Jyväskylä, Anna moved into neuroscientific research and was granted ERASMUS funding for an internship in the neurocognition of music with Daniela Sammler at the Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany.

In 2019, Anna started PhD training at the Max Planck Institute for Empirical Aesthetics, Frankfurt am Main, Germany, subsequently enrolling at the Faculty of Psychology and Neuroscience at Maastricht University, The Netherlands. In Frankfurt, Anna conducted concert research supervised by Melanie Wald-Fuhrmann and Lauren Fink. In Maastricht, Anna was supervised by Sonja Kotz for the experimental and methodological part of her research. During this time, Anna also taught empirical methods in music research at the Hochschule für Musik in Karlsruhe, Germany. She has also taken on the role of PhD representative for the MPI for Empirical Aesthetics in 2021-2022 and been a member of the Max Planck PhDNet Open Science working group since 2021. As a musician, Anna enjoys playing piano, violin, and sings. Performance highlights include playing with the Oxfordshire Youth Baroque Strings (Oxybaroxy), singing with Coro Piccolo in Leipzig, and performing as the soloist in *Bird Concerto with Pianosong* with The Chimera Ensemble.

Appendix

List of Publications

Peer-reviewed publications

Czepiel, A., Fink, L. K., Seibert, C., Scharinger, M., & Kotz, S. A. (2023). Aesthetic and physiological effects of naturalistic multimodal music listening. *Cognition*, *239*, 105537. https://doi.org/10.1016/j.cognition.2023.105537

Gibbs, H. J., **Czepiel, A.**, & Egermann, H. (2023). Physiological synchrony and shared flow state in Javanese gamelan: Positively associated while improvising, but not for traditional performance. *Frontiers in Psychology*, *14*, 1214505. https://doi.org/10.3389/fpsyg.2023.1214505

Czepiel, A., Fink, L. K., Fink, L. T., Wald-Fuhrmann, M., Tröndle, M., & Merrill, J. (2021). Synchrony in the periphery: Inter-subject correlation of physiological responses during live music concerts. *Scientific Reports*, *11*(1), 1-16. https://doi.org/10.1038/s41598-021-00492-3

Merrill, J., **Czepiel, A**., Fink, L. T., Toelle, J., & Wald-Fuhrmann, M. (2021). The aesthetic experience of live concerts: Self-reports and psychophysiology. *Psychology of Aesthetics, Creativity, and the Arts*, *17*(2), 134–151. https://doi.org/10.1037/aca0000390

Tschacher, W., Greenwood, S., Egermann, H., Wald-Fuhrmann, M., **Czepiel, A.**, Tröndle, M., & Meier, D. (2021). Physiological synchrony in audiences of live concerts. *Psychology of Aesthetics, Creativity, and the Arts.* https://doi.org/10.1037/aca0000431

Wald-Fuhrmann, M., Egermann, H., **Czepiel, A.**, O'Neill, K., Weining, C., Meier, D., Tschacher, W., Uhde, F., Toelle, J., & Tröndle, M. (2021). Music Listening in Classical Concerts: Theory, Literature Review, and Research Program. *Frontiers in Psychology*, 12. https://www.frontiersin.org/article/10.3389/fpsyg.2021.638783

Preprint articles

Merrill, J., Ackermann, T. I., & **Czepiel, A.** (2023). The negative power of music: Effects of disliked music on psychophysiology. *Psyrxiv.*

Manuscripts in preparation

Criscuolo, A., **Czepiel, A**., Schwartze, M., Kotz, S. A., (In preparation). Body and brain physiological waves from health to pathology.

Czepiel, A., Fink, L. K., Seibert, C. Scharinger, M., Wald-Fuhrmann, M., Kotz, S. A., (In preparation). Cardiorespiratory synchrony in live music concerts: is it best predicted by stimulus, closeness to other people or aesthetic experience?

Czepiel, A., Sammler, D., (In preparation). Motor simulation and context information facilitate anticipation in expert pianists.

Conference contributions

Czepiel, A., Fink, L. K., Seibert, C., Scharinger, M., Wald-Fuhrmann, M., & Kotz, S. A. (2023, July). *Cardiorespiratory synchrony to music and across audience members in concerts.* Talk presented at 17th International Conference on Music Perception and Cognition (ICMPC 17), Tokyo, Japan.

Czepiel, A., Fink, L. K., Seibert, C., Scharinger, M., Wald-Fuhrmann, M., & Kotz, S. A. (2023, June). *Cardiorespiratory synchrony during live music concerts.* Talk presented at Cognitive Neuroscience on Body-Brain: Waves. Salerno, Italy.

Czepiel, A., Sammler, D., (2023, June). *Motor simulation and context information facilitate action and anticipation in expert pianists.* Poster presented at the Symposium on System Neuroscience: Audiomotor Integration in Cognition: Comparative Neurobiology in Human and Non-human Primates, Querétaro, Mexico.

Czepiel, A., Fink, L. K., Seibert, C., Scharinger, M., & Kotz, S. A. (2022, May). *Physiological correlates of aesthetic and naturalistic music concert experience.* Poster presented at the International conference of Cognitive Neuroscience (ICON), 2022. Helsinki, Finland.

Czepiel, A., Fink, L. K., Seibert, C., Scharinger, M. Kotz, S. A., (2021, July). *Multimodality of music listening: how live versus recorded versions of piano music influence aesthetic, physiological and neural responses in a concert setting.* Poster at International Conference for Music Perception and Cognition 16 and European Society for Cognition of Music 11 (ICMPC16-ESCOM11), 2021. (Virtual).

Czepiel, A., Fink, L. K., Fink, L. T., Wald-Fuhrmann, M., Tröndle, M., *&* Merrill, J. (2021, May). *Inter-subject correlation of physiological responses during live musical performance.* Poster presented at Neurosciences and Music VII, 2021. (Virtual).

Czepiel, A., Fink, L. K., Fink, L. T., Wald-Fuhrmann, M., Tröndle, M*., &* Merrill, J. (2021, March). *Inter-subject correlation of physiological responses during live musical performance.* Poster presented at Cognitive Neuroscience Society Annual Meeting 2021. (Virtual).

Czepiel, A., Fink, L. K., Seibert, C., Scharinger, M. (2021, March). *Multimodality of music listening: audio-visual influences on aesthetic experience during piano music.* Presentation at Tagung Experimentell Arbeitender Psychologen (TeaP), 2021. (Virtual).

Czepiel, A., Fink, L. K., Fink, L. T., Wald-Fuhrmann, M., Tröndle, M*., &* Merrill, J. (2020, September). *Physiological responses to Classical music in a concert: inter-subject correlation analysis.* Poster presented at 36 Jahrestagung der Deutschen Gesellschaft für Musikpsychologie e.V. (DGM), 2020. (Virtual).

Czepiel, A., (2020, March). *Investigating the aesthetic experience in live concerts with continuous measures: how psychophysiology reflects temporal development of music.* Talk presented at Arbeitstagung der Fachgruppe Systematische Musikwissenschaft, 2020. Kassel, Germany

Czepiel, A., Egermann, H., Toelle, J. O'Neill, K. Weining, C., (2019, July). *Measuring the audience experience – how can we capture audience experience with quantitative and qualitative methods?* Panel discussion. Audience Research in the Arts, 2019. Sheffield, United Kingdom

Czepiel, A., Sammler, D., (2019, June). *Go with the flow: continuous kinematic information facilitates action planning of performing chord sequences in expert pianists.* Poster presented at 9th IMPRS NeuroCom Summer School, 2019. Leipzig, Germany

Czepiel, A., Sammler, D., (2019, March). *The role of continuity and action simulation in detecting action errors in pianists: an ERP study.* Poster presented at Mind Brain Body Symposium, 2019. Berlin, Germany

Czepiel, A., (2019, February). The Role of Harmonic Flow in Processing Syntax when Imitating Musical Actions: an ERP Study, Spring School 2019: *Language and Music in Cognition*, University of Cologne, Cologne, Germany

Czepiel, A. Luck., G, (2018, July). 'Importance of felt mood and emotion for expressive and technical movement characteristics in pianists.' 15th International Conference on Music Perception and Cognition and 10th Triennial Conference of the European Society for the Cognitive Sciences of Music, University of Graz, Graz, Austria

Czepiel, A., Allingham, E., Oudyck, K., Zamundio, A., Saari, P. (2017, September). 'Musicians' timbral adjustments in response to emotional cues in musical accompaniments.' SysMus, Queen Mary University of London, United Kingdom

Czepiel, A., Hollingworth, R., Egermann, H., (2017, June). 'Conscious and subconscious body movement: an exploratory study into technical and expressive gestures in piano players.' 25th Anniversary Conference of the European Society for Cognitive Science in Music, University of Ghent, Ghent, Belgium.

Czepiel, A., Hollingworth, R., Egermann, H., (2016, June). 'Conscious and subconscious body movement: a study into technical and expressive gestures in pianists.' Music Education and Psychology Student Conference, University of York, York, United Kingdom

Presentations and outreach

Czepiel, A. M., (2022, November). Talk on Open Science in the Max Planck Institute, on behalf of the PhDNet Open Science Working Group. At Max Planck Institute for Security and Privacy in Bochum, as part of the *Spark(l)ing Science* talk series.

Czepiel, A. M., (2022, October). Talk on Open Science in the Max Planck Institute, on behalf of the PhDNet Open Science Working Group. At Max Planck PhDNet General Meeting, Cologne.

Czepiel, A. M., (2022, October). Talk and workshop at Donders institute, Nijmegen, 'Mobile measurement of body-brain signals in naturalistic settings.'

→ Workshop available on the Open Science Framework: https://osf.io/a7yfn/

Czepiel, A. M., (2022, May). *Out of the Lab and into the Wild: Measuring music experience in naturalistic settings.* Keynote at Synapsium Dondrite, Donders Institute, Nijmengen.

Supervision

Supervision (Korreferent) of Till Gerneth, Master's thesis, (2022) 'The Emotional Influence of Headphone and Speaker Reproduction in Music Listening', Hochschule Darmstadt

Appendix

Translations into Dutch

Title:

Het bestuderen van muziek bij het luisteren naar concerten: esthetische ervaringen en perifere fysiologische reacties

Impact paragraph

Het luisteren naar live muziek is sociaal en cultureel gezien een belangrijke menselijke activiteit. We kiezen ervoor om tijd en geld in deze activiteitte investeren, omdat het ons leven emotioneel, intellectueel en sociaal verrijkt. Voor 2020 nam het bezoek aan live muziekconcerten in landen als het Verenigd Koninkrijk (Art Council UK, 2014/2015), Duitsland (Deutscher Bühnenverein, 2022) en Nederland (Grace, 2020) over het algemeen toe. Het belang van livemuziek werd benadrukt door de lockdowns tijdens de COVID-19 pandemie, waar muziek een krachtig middel was voor sociale verbondenheid in een schijnbaar donkere tijd (Fink et al., 2021). Het stoppen van de financiering van nationaal belangrijke culturele instellingen zoals de English National Opera (Connolly, 2022) en de BBC (British Broadcasting Corporation) Singers (Kelly, 2023) heeft tot hevige kritiek geleid van musici, wetenschappers en de samenleving.

Deze voorbeelden laten de sociale en culturele waarden van live muziek zien. Daarom was het doel van het huidige onderzoek om ons begrip van live muziek luisterervaringen te vergroten, met name in westerse klassieke muziekconcerten. Om de concertervaring te beoordelen, werd in het huidige onderzoek gebruik gemaakt van zelfrapportage- enquêtes en perifere fysiologische metingen. Deze metingen bestonden uit reacties van het vecht-en-vluchtsysteem (die meten hoe aandachtig iemand luistert en hoe emotioneel opwindend de muziek is) en de activatie van gezichtsspieren (fronzen en glimlachen). Hoe onze bevindingen verband houden met welzijn en toekomstig onderzoek onderbouwen, wordt hieronder besproken.

Uit eerdere studies weten we dat het bijwonen van en ook het optreden tijdens live muziek evenementen een positief effect kan hebben op het welzijn. De positieve esthetische ervaringen die muziek kan oproepen zijn sterker bij live muziek (Gabrielsson & Wik, 2003; Lamont, 2011). Een belangrijke bevinding van het huidige proefschrift was dat concerten inderdaad een hoge waardering en positieve emoties oproepen, ondanks incidentele negatieve emoties. Verder werd deze positieve ervaring versterkt bij live, audiovisuele muziekuitvoeringen, in vergelijking met alleen audio-opnames. Dit suggereert dat bepaalde componenten van concertprogramma's samen met hun visuele aspecten belangrijke elementen kunnen zijn die positieve ervaringen oproepen. Deze positieve ervaringen komen op hun beurt waarschijnlijk het algehele welzijn ten goede. Door de (synchronie van) fysiologische reacties te beoordelen, heeft dit proefschrift ook potentiële mechanismen aan het licht gebracht van hoe dergelijke positieve gevoelens worden opgeroepen, namelijk door betrokkenheid, spiegeling van emotie en motorische nabootsing. Hoewel het huidige onderzoek werd uitgevoerd met over het algemeen gezonde deelnemers, zou een dergelijk begrip van positieve ervaringen in de toekomst kunnen worden uitgebreid om na te gaan of deze positieve invloed op het welzijn ook gunstig kan zijn voor de algemene (lichamelijke en geestelijke) gezondheid.

Concerten hebben belangrijke sociale voordelen. Sociale verbinding is inderdaad een evolutionaire verklaring van muziek (Kotz et al., 2018; Savage et al., 2021). Live muziekoptredens creëren belangrijke sociale ervaringen (Dearn & Price, 2016; Packer & Ballantyne, 2011). Ze zorgen ook voor een gevoel van saamhorigheid en gemeenschapsgevoel (Pitts, 2004). De collectieve ervaring werd in het huidige proefschrift onderzocht door het berekenen van synchronie in tijd en fase van de fyiosologische reacties in het publiek (hoofdstuk 3 en 5). Uit deze resultaten blijkt dat muziek de fysiologische reacties van de toeschouwers synchroniseert, hoewel de tijdsynchronisatiemetingen robuuster waren dan de fasesynchronisatiemetingen. Muziek synchroniseerde bijvoorbeeld de versnelling en vertraging van de hartslag van de toeschouwers, maar niet hun eigenlijke hartslagen. Eén reden voor het gebrek aan fasesynchronie wordt verklaard door de gebruikelijke beperkte fysieke beweging en interactie van het publiek bij westerse klassieke concerten. Bij andere genres, zoals pop en rock, is meer interactie door samen te dansen en mee te zingen met de muziek niet alleen toegestaan, maar ook een positief onderdeel van de ervaring. In deze genres is de fasesynchronie misschien relevanter. Hoewel dit proefschrift het sociale aspect niet rechtstreeks onderzocht, zouden toekomstige studies een soortgelijk paradigma en analysestappenplan kunnen toepassen om de mechanismen van sociale verbinding tijdens het luisteren naar muziek verder te evalueren en te begrijpen. Het synchronie analyse stappenplankanworden toegepast op andere soorten signalen zoals hersensignalen, eye tracking, en maagritmes, om verdere informatie te verkrijgen over hersen-lichaam interacties. Dergelijke metingen kunnen niet alleen worden toegepast ophet luisteren en uitvoeren van muziek, maar ook op andere situaties, zoals onderwijs (bijvoorbeeld synchronie tussen leraar en leerlingen) en sociale activiteiten zoals gesprekken, dans en teamsport.

De resultaten van dit onderzoek kunnen informatie opleveren voor degenen die concerten plannen en organiseren. Westerse klassieke concerten worden over het algemeen gezien als zeer formele evenementen, die de opkomst niet altijd bevorderen. In mijn eigen werk als concertmanager (zie bijlage, C.V.) heb ik kunnen ervaren hoeveel moeite concertorganisatoren doen om hun publiek te diversifiëren door een zorgvuldige planning van het concertprogramma, door educatieve aspecten op te nemenen door de interactie met het publiek te vergroten. Dit proefschrift breidt ons empirisch begrip uit van aspecten die concerten aantrekkelijker kunnen maken, bijvoorbeeld een evenwichtig programma van bekende muziek, gemengd met nieuwe, mogelijk enigszins uitdagende muziek. Dit en toekomstig onderzoek kan licht werpen op mogelijke aandachtspunten voor de marketing van concerten voor een breder publiek. Dergelijke investeringen zijn van vitaal belang, omdat concerten niet alleen in cultureel, maar ook in economisch opzicht de samenleving ten goede komen. Succesvolle concerten en muziekfestivals hebben de zichtbaarheid en het toerisme in kleinere steden en plaatsen vergroot, bijvoorbeeld Darmstadt (bekend om de Nieuwe Muziek), Bayreuth (bekend om het wereldberoemde Wagnerfestival) en Aldeburgh (bekend om zijn innovatieve muziekfestival). Meer concertbezoekers zouden de kans vergroten dat investeringen worden gedaan om muziek op grotere schaal te promoten. Dit zou vervolgens de financiering kunnen ondersteunen die nodig is om muziek betaalbaarder te maken, wat op zijn beurt de samenleving in het algemeen ten goede komt.

Kortom, concertonderzoek en concertbezoek zijn dus niet alleen waardevol voor het positieve welzijn en het sociale nut van individuen, maar hebben ook een aanzienlijke maatschappelijke en culturele waarde.

Om dit onderzoek zo toegankelijk mogelijk te maken zijn of worden de onderzoeken in hoofdstuk 2-5 gepubliceerd in preprints (voorafgaand aan paper acceptatie) en open access tijdschriften. Dit proefschrift draagt ook bij aan Open Science, d.w.z. praktijken die wetenschap beschikbaar, toegankelijk, herbruikbaar en transparant maken. De gegevens van het pianoconcert en de analysecode zijn geüpload naar het Open Science Framework. Dit is momenteel privé zolang het paper in review is, maar het zal worden openbaar gemaakt zodra het artikel is geaccepteerd. Het onderzoek in dit proefschrift is gepresenteerd als een keynote presentatie en als posters en lezingen op andere internationale wetenschappelijke conferenties. Deze presentaties dragen bij aan het groeiende wetenschappelijke veld van het begrijpen van het lichaam en de hersenen (Criscuolo et al., 2022) en aan de vooruitgang van de neurowetenschappen in reële situaties (Tervaniemi, 2023).

Appendix

Acknowledgements

I would like to thank so many people that made this PhD not only possible, but also an incredible scientific and personal journey. I would like to thank Melanie Wald-Fuhrmann, who gave me the opportunity and financial resources to do my PhD at the Max Planck Institute for Empirical Aesthetics; I am very grateful for allowing me the freedom to pursue my own scientific interests. Thank you for your expertise and in helping me think more broadly in a range of disciplines.

I wish to also thank Sonja A. Kotz. I am so fortunate that you accepted me as a PhD at Maastricht University and I am deeply grateful that you dedicated so much time into supervising and supporting me. It has been incredible to learn from you and I want to thank you for your scientific guidance, wisdom, and motivation. Your dedicated and thoughtful feedback has not only helped improve and shape my thesis, but has also pushed me to be a better scientist. Thank you for being a truly inspiring role-model.

I am also very thankful to Lauren K. Fink. I am so grateful that you helped me develop my skills as an interdisciplinary researcher; in particular for teaching me to always question myself and check my data and all analysis steps visually. Thank you for all your expertise, professional support, and friendship!

I am particularly grateful to many colleagues that have been part of the concert research projects. A huge thanks go to Julia Merrill, who helped me so much when I first arrived at the MPI-AE. Thanks for your enthusiasm, discussions, and the scientific opportunities! Thanks also go to Jutta Toelle, Mathias Scharinger, and Christoph Seibert. I am grateful that I was involved for some time with Experimental Concert Research team. Thanks go to Christian Weining, Deborah Meier, and especially Mimi O'Neill – the support of you PhDs was so invaluable to me and it was great fun setting up and down the concert experiments with you!

A special thanks goes to the Daniela Sammler, who already helped me begin my research journey in Germany and at the Max Planck Institute during my internship in 2018. I am so grateful for your incredible support and encouragement throughout the years. I am particularly thankful that I could continue joining the Neurocognition of Music and Language meetings! It was a great opportunity to learn and have insightful discussions with the groups. In particular, thanks go to Natalie Kohler, Kasia Gugnowska, Hannah Ringer, Seung-Goo, and Seung-Cheol who were supportive at various stages and happy to help with general methods questions.

I would also like to say a huge thank you to my colleagues in the BAND-Lab at Maastricht University who made me always feel very welcome in our Monday meetings when I joined online. In particular, I am very grateful to Antonio Criscuolo and Hanna Honcamp, who organised a wonderful schedule of insightful talks, invited speakers, workshops and journal clubs. It was wonderful to be a part of this team and I learnt a lot from our discussions.

I am grateful that I could learn so much from others in the scientific field. I am also grateful to members of the music department at MPI-AE, to whom where I could present and gain valuable advice. I wish to also acknowledge the support of Myriam Mieles and Silke Yalin, who helped me through various help regarding administrative side of working as an international in Germany. Thanks also go to Daniel Shanahan, who was a great mentor during my time at the Women in MIR. I am so grateful also to members of the MPG PhDNet Open Science working group – I have learnt so much being a part of this and has inspired me to continue open science practices in my future endeavours.

My PhD life has been also enriched by so many wonderful people at the Max Planck Institute. I am especially thankful for the PhD support network that we have at MPI-AE. Special thanks go to the Monika Tschense, who shared in so many of my ups and downs throughout the thesis; thanks for all the lovely choir singing and café PhD writing sessions. Thanks to Birte Thissen and Victoria Poehls (missing our Fish Fridays and Aroma sessions!). Thanks go Christina Lubinus for our coffee dates, PhD support, and being a great conference buddy! To Iris Mencke, who gave me so much wise advice on music and neuroaesthetics – thank also for all the lovely sushi dates! A huge thanks to Lea Fink, for all your inspiring music theory chats and looking forward to more music making in the future! Thanks also go to Kanthida van Welzen who did an incredible job with the PhD representative when I had too much on my place and bringing the community together after the COVID-19. I also thank Kanthida and Pol van Rijn for helping translate some of the text into Dutch :)

I am very grateful also to several friends who reminded me take time away from the desk. To Lara Pearson, thanks for all the delicious dinners, noodles, and birding walks! To my flatmate Ava Kiai, for our fun violin duets and fab film nights! To Peter Harrison, Manu Anglada-Tort; thank you for all the wonderful and fun evenings of music jams, food, drinks and for being wonderful conference buddies! Thanks to my ultimate babes Emma Allingham and Alvaro Chang, whose zoom calls about PhD life and psychology methods were so valuable to me - looking forward to making more music and research with you in the future. Thank you to my dear long-term friends who accompanied me throughout this time: Athene Broad, Holly Gurney and the fab FAME gals, Miri Nohl, Elinor Bishop, and Fiona Mayho; Cameron Butler, and Jamie Herzig. Thanks for providing me with happiness and sunshine!

I work well when I am listening to something; therefore, I feel I can't write acknowledgments without acknowledging some of the people who were accompanying me indirectly during my PhD work. Thank you, Taylor Swift, whose music I listened to almost exclusively for so many of the months of my PhD. Thanks go to some of the composers who kept me most company during the writing: Gustav Mahler, Richard Strauss, and Jean Sibelius. Thank go to the wonderful people at BBC Radio 3, in particular to Andrew Macgregor 'Record Review', for introducing amazing new records that would motivate my writing. Thanks also go to Iliza Shlesinger (and co-host Emily Higgins), whose inspiring podcast made me laugh and kept me sane during data pre-processing session.

To my partner(-in-PhD-crime), who supports and puts up with all the hard work I do, thank you Lucas Rickert: for giving me so much motivation and kind words over the course of the PhD. Thank you for all the pity parties and all celebrations – no matter how big or small the reason. Last and by no means least, I am HUGELY indebted to my wonderful family. Thanks so much to my inspirational grandmas, my brother Mikey and my sister Magda (who made the BEAUTIFUL cover for my thesis!), for bringing me joy and encouragement. Most importantly, thank you so much to my incredible parents: for always being there, for believing in me, and for being the best proof-readers in the world. I am so grateful for all the never-ending faith and inspiration, encouragement to always make the best out of anything, and for all the unconditional love. Thank you for being my biggest fans!